

Proyección de poblaciones carcelarias en Colombia

SERGIO DAVID SOLANO BEJARANO
INGENIERO INDUSTRIAL



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICA
BOGOTÁ, D.C.
ABRIL DE 2017

Proyección de poblaciones carcelarias en Colombia

SERGIO DAVID SOLANO BEJARANO
INGENIERO INDUSTRIAL

DISERTACIÓN PRESENTADA PARA OPTAR AL TÍTULO DE
MASTER EN CIENCIAS - ESTADÍSTICA

ADVISOR
B. PIEDAD URDINOLA CONTRERAS, PH.D.
DOCTOR EN DEMOGRAFÍA

RESEARCH LINE
DEMOGRAFÍA



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICA
BOGOTÁ, D.C.
ABRIL DE 2017

Title in English

Prison populations projections for Colombia

Título en español

Proyección de poblaciones carcelarias en Colombia

Abstract: Not there yet!

Resumen: Se presentan proyecciones de población carcelaria cuando no se conocen las tasas de ingreso al sistema carcelario, el tiempo de juicio, ni la duración de la pena. Para estimar estas variables no observadas se utilizan modelos estado espacio multivariados, tomando como base la población sindicada y condenada mensual

Keywords: Prison populations, time series, SARIMA processes, State Space Models

Palabras clave: Poblaciones carcelarias, series de tiempo, procesos SARIMA, Modelos Estado Espacio

Acceptation Note

Thesis Work

Aprobado

“Meritoria o Laureada mention”

Jury

Jurado uno

Jury

Jurado dos

Advisor

B. Piedad Urdinola

Bogotá, D.C., Abril 19 de 2017

Dedicado a

El problema es la codificación de los archivos, que cambia.

Agradecimientos

Tíldes corregidas.

Contents

Contents	I
List of Tables	III
List of Figures	IV
Introducción	V
.1 Objetivos	VI
.1.1 Objetivo General	VI
.1.2 Objetivos Específicos	VI
1. Antecedentes teóricos	1
1.1 Proyecciones de población	1
1.1.1 Proyección de poblaciones pequeñas	2
1.1.2 Aplicaciones nacionales e internacionales	2
1.2 Series de tiempo	3
1.2.1 Modelos ARIMA, SARIMA	4
1.2.2 Modelos Estado-Espacio	4
2. La población carcelaria en Colombia 1991 - 2017	5
2.1 Análisis exploratorio	5
2.2 El sistema penitenciario en Colombia	6
2.3 Identificación del modelo	6
3. Modelos SARIMA	8
3.1 Marco teórico	8
3.2 Identificación del modelo	8
3.3 Proyecciones 2017 - 2020	8

3.4 Conclusiones	8
4. Modelos estado-espacio	9
4.1 Marco teórico	9
4.2 Identificación del modelo	9
4.3 Simulación Monte-Carlo	9
4.4 Estimación de parámetros	9
4.5 Proyecciones 2017 - 2020	9
4.6 Conclusiones	9
Conclusiones	10
Trabajo futuro	11
Bibliography	12

List of Tables

List of Figures

2.1	Población privada de la libertad 1991 - 2017	5
2.2	Tasa de encarcelamiento según género 1991 - 2017	6
2.3	Población carcelaria por situación jurídica	7

Introducción

La población carcelaria es el grupo poblacional que se encuentra privado de la libertad en una institución penitenciaria. Dentro de esta población, a aquellos a la espera de juicio se les denomina sindicados, y condenados a quienes están cumpliendo una sentencia. En Colombia el INPEC (Instituto Nacional Penitenciario y Carcelario), que "es la institución pública administradora del sistema penitenciario y carcelario del país" [7], publica periódicamente la serie histórica de población carcelaria, que da cuenta de la población al cierre de cada mes desde 1991, separada por situación judicial (sindicados, condenados) y género.

De esta serie de población El CONPES 3828 anota que "el promedio anual de crecimiento de la Población Privada de la Libertad (PPL) entre 1993 y 2014 fue de 9,43% para la población condenada y 4,91% para la sindicada" [5]; ambas tasas están muy por encima del 1,18% anual, crecimiento poblacional nacional estimado por el DANE para el quinquenio 2005-2010 [4]. El aumento por encima de la población nacional sugiere dinámicas de crecimiento diferentes al resto de la población.

De otra parte el mismo documento observa que "en el periodo 1993-2014 la proporción de cupos habilitados creció en 173,59%, mientras la población privada de la libertad (PPL) creció en un 315,39%." [5]. El crecimiento acelerado de la población carcelaria, por encima de la oferta de cupos implica un incremento del hacinamiento, lo que a su vez podría incidir en problemas de salud y de orden al interior de las instituciones carcelarias.

El CONPES 3828 de 2015 realiza, además, proyecciones de la población carcelaria a través de un modelo de crecimiento exponencial con tasa de crecimiento de 7,23% anual. Este tipo de modelos, aunque ampliamente usados en proyecciones de población, no resulta conveniente cuando la tasa de crecimiento no es estable a través del tiempo y depende de variables exógenas como: la población nacional, su estructura etaria, las características del sistema penitenciario y del sistema judicial.

En este contexto resulta necesario contar con proyecciones de la población carcelaria en diferentes escenarios, para determinar la necesidad de cupos en el corto y mediano plazo. Tales proyecciones deben tener en cuenta las dinámicas propias del sistema carcelario y deben permitir estimar el impacto de modificaciones a la política criminal y carcelaria, de forma que permitan una mejor planeación de la cantidad y/o tamaño de los centros carcelarios. Este problema se ha abordado anteriormente desde el estudio de poblaciones pequeñas. [11]

Las poblaciones que se estiman y proyectan a un nivel menor que el nacional se conocen como poblaciones pequeñas (ciudad, departamento, etc...). Dentro de las poblaciones

pequeñas se encuentran poblaciones especiales que "se localizan en un área especial por una acción legislativa o administrativa ". [11]. La población carcelaria es un tipo de población especial cuyos componentes: nacimientos y defunciones no necesariamente siguen los patrones biológicos que dan vida al modelo exponencial y requieren el uso de técnicas diferentes que permitan un mejor ajuste.

Las series de población carcelaria disponibles para Colombia son mediciones de la población al cierre de mes, agrupada por situación judicial y género. Las proyecciones de población suelen incluir las tasas de nacimiento, migración y mortalidad. En el caso de las poblaciones carcelarias, los equivalentes serían la tasa de ingreso al sistema, la tasa a la que son juzgados, la mortalidad y la duración de las penas, que determina la tasa de salida de los condenados. En el caso de Colombia estos datos no son libres, por lo tanto el modelo debe permitir estimar las tasas no observadas. Adicionalmente, estas tasas no son constantes a través del tiempo, pues dependen de la estructura del sistema judicial y del sistema carcelario.

Este trabajo busca ofrecer proyecciones de población carcelaria de corto y mediano plazo, comparando los modelos de: series de tiempo, modelos estado espacio y métodos de proyección de poblaciones pequeñas.

.1 Objetivos

.1.1 Objetivo General

Comparar el ajuste de los métodos clásicos de proyección de poblaciones pequeñas (ratio-correlation, Censal-ratio methods) con el ajuste de series de Tiempo ARMAX y de modelos Estado-Espacio, para la población carcelaria de Colombia 1991-2016.

.1.2 Objetivos Específicos

- Proyectar la población carcelaria usando los métodos de ratio-correlation y censal-ratio. (0%)
- Proyectar el comportamiento de la población carcelaria a través de series de tiempo ARIMA y ARIMAX, suponiendo tasas de transición estables. (PROCESO) ()
- Aplicar modelos estado espacio para estimar las tasas de transición, cuando: son no observadas, varían en el tiempo y tienen correlación con variables exógenas. (PROCESO)
- Comparar el ajuste de los tres métodos a través de simulación Montecarlo. (Nulo)
- Aplicar los métodos descritos en los objetivos anteriores, a los datos de población carcelaria colombiana 1991-2016. ()

CHAPTER 1

Antecedentes teóricos

En 2016 Colombia ocupa el puesto catorce entre doscientos cincuenta y un países por el tamaño de su población carcelaria (120 914 hbts.) y el cincuenta y uno según la tasa de encarcelamiento (240 por cada 100.000 hbts). Tasa que pasó de 51,5 en el año 2000 a 240 por cada 100.000 hbts en 2016. Con una ocupación del 154% de las plazas disponibles, resulta relevante contar con proyecciones de la población carcelaria en el corto, mediano y largo plazo. [6]

1.1 Proyecciones de población

"Una estimación poblacional consiste en determinar el tamaño o las características de una población, para el momento actual o para uno anterior, en ausencia de información. Cuando se realizan un conjunto de supuestos sobre el comportamiento de los vitales hacia el futuro, hablamos de proyección, y cuando se escoge un escenario como el más probable, hablamos de pronóstico"[11].

Para incluir la incertidumbre en las proyecciones de población Lee enumera los siguientes métodos [9]:

- El enfoque de escenarios alto, medio y bajo: Asume comportamientos fijos para la fertilidad, la mortalidad y las migraciones durante el periodo de proyección, basado en algunos supuestos [9].
- Análisis estocásticos
 - Análisis ex-post: consiste en evaluar el error de pronóstico en proyecciones anteriores y aplicarlo a las nuevas proyecciones [9].
 - Simulación estocástica: Permite hacer proyecciones de población, al asignar una distribución de probabilidad a las tasas vitales (mortalidad, natalidad, migraciones) [9].
 - Modelos estocásticos de la tasa de crecimiento: Consiste en estimar la tasa de crecimiento del total de la población; aunque permite estimar intervalos de confianza, no permite separar la proyección de las tasas vitales, ni de las franjas etarias [9].

- Matrices de Leslie con modelos estimados para las tasas vitales: Al usar matrices de Leslie se estima la población por rangos etarios para un instante i , y se calcula la población en el instante $i + 1$ aplicando la natalidad y la mortalidad proyectadas para el periodo i . Puesto que las series de población carcelaria no se publican separadas por edad, no se abordará esta técnica.

1.1.1 Proyección de poblaciones pequeñas

La proyección de áreas pequeñas es entendida como la proyección a un nivel geográfico menor al nacional. Estas proyecciones pueden incluir, departamentos, ciudades o poblaciones especiales [11]. "Una población especial es un grupo poblacional que se encuentra restringido a un área por una medida administrativa o legislativa. Dentro de los grupos usualmente considerados se encuentran las prisiones, universidades, hospitales e instituciones militares". Este tipo de población puede tener una estructura etaria y de sexo, y unos vitales diferentes al resto de la población; además no suelen envejecer en el mismo lugar, lo que permite mantener una estructura etaria que no varía a través del tiempo. [11].

1.1.2 Aplicaciones nacionales e internacionales

Las proyecciones de poblaciones carcelarias oficiales analizadas corresponden, en buena parte, a los métodos expuestos en los capítulos anteriores: Proyecciones por escenarios, proyección de la tasa de crecimiento, modelos ARIMA para las tasas de ingreso y salida.

En Colombia (CONPES 3828) se proyectó la población carcelaria usando la tasa media de crecimiento anual (1993-2014) [5]. Esta proyección no tiene en cuenta la incertidumbre asociada con las variaciones aleatorias en las tasas, ni las asociadas a cambios estructurales.

El Reino Unido hasta 2015 realizaba una proyección por escenarios (alto, medio y bajo), año en el cual cambió a un modelo de proyección de la media y su incertidumbre. La incertidumbre se incluyó a través de un análisis ex-post, de la desviación de la proyección en años anteriores [8].

El departamento de Justicia de los Estados Unidos realizó estimaciones de la población carcelaria por estado para el periodo 2013-2014. Las estimaciones parten del censo de prisiones 1993-2014. Estas proyecciones se puede enmarcar dentro de las proyecciones de áreas pequeñas [10].

El bureau de estadísticas e investigación del crimen en Australia proyecta las tasas de arresto y sentencia usando modelos ARIMA; a partir de estas tasas proyecta la población carcelaria. Estas proyecciones incluyen un periodo de validación de tres años. Los resultados mostraban que la serie real se encuentra dentro de los intervalos de confianza de la proyección, cercano a la proyección de la media [13].

Blummstein desarrolla un método de proyección basado en los componentes demográficos, tasas específicas de arresto por delito y reincidencias, a partir de estos datos proyecta el tamaño y la composición de las poblaciones [1].

Con los datos libres disponibles en Colombia no se podría utilizar el enfoque ARIMA ni el método de Blummstein, pues las tasas de encarcelamiento y sentencia no se publican.

La tesis busca proponer un método de proyección, para situaciones donde no se cuenta con el registro de los vitales o su equivalente en la población analizada.

1.2 Series de tiempo

Una serie de tiempo es un conjunto de observaciones x_t asociadas a un instante de tiempo t . Es usual referirse como series de tiempo, tanto a las realizaciones x_t como a las variables aleatorias X_t que las generan.[2]

En una regresión lineal clásica, una variable Y es explicada o predicha en función de una variable X . La diferencia entre el valor observado y_i y el valor predicho se suponen provenientes de un proceso aleatorio con media cero. [3]

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Donde $\epsilon_1, \epsilon_2, \epsilon_3$ son independientes e idénticamente distribuidas (i.i.d.). Es posible representar una serie de tiempo en esta forma, tomando como variable explicativa el tiempo. Sin embargo, el análisis de series de tiempo se ha desarrollado como un área particular de la estadística, pues es este supuesto no suele cumplirse, enfocándose particularmente en variables aleatorias dependientes, que pueden ser representadas en la forma:

$$Y_i = \beta_0 + \beta_1 Y_1 + \beta_2 Y_2 + \dots + \beta_{i-1} Y_{i-1} + \epsilon_i$$

Estacionaridad. Estacionaridad debil.

Cómo se puede medir esto? con una función de autocorrelación, que mide la relación entre los valores del proceso, según su cercanía.

Time series is therefore primarily concerned with dependent random variables and the analysis and utilization of this dependence.) A simpler approach to forecasting X_{n+h} is to look for the linear combination, $X_{n+h} = a_0 + a_1 X_n + \dots + a_n X_1$ which minimizes the expected squared error $E(X_{n+h} - \hat{X}_{n+h})^2$. This is a much simpler problem, the solution of which depends only on the expected values $E X_i$ and $E X_i X_j$, $i, j = 1, 2, \dots$. Moreover if the joint distribution of (X_1, X_2, \dots, X_k) is multivariate normal for every positive integer k then this best linear forecast

Las series de tiempo tienen una aplicación bastante limitada, pues requieren muchas (?) observaciones, espaciadas uniformemente. Cuando la correlación se presenta en lags uniformemente espaciados por ejemplo (12,24,36) se denomina una serie estacional.

Para medir que tan bien ajusta el modelo analizamos el comportamiento del error de estimación. Es usual suponer que el error tiene varianza constante (homocedasticidad), es independiente e idénticamente distribuido, usualmente con distribución normal.

Es posible que el modelo tenga tendencia. Cuando la tendencia es lineal tomamos la diferencia al lag 1 (si queremos desestacionalizar, tomamos la diferencia al lag que nos sirve para desestacionalizar) y evaluamos la autocorrelación en esta nueva serie. (Expandir con libro del curso de Datacamp)

Cuando la varianza incrementa con el tiempo podemos asumir que el modelo crece de manera exponencial (un incremento porcentual fijo). Para manejar estas series de manera sencilla tomamos la logaritmo natural de la diferencia entre pasos. (Expandir con Tsay)

1.2.1 Modelos ARIMA, SARIMA

El análisis de las series tiempo de poblaciones carcelarias de tiempo ARIMA y ARMAX se tomará de [12].

Los procesos ARMA son procesos aleatorios de la forma

$$Y_t = \gamma Y_{t-1} + \gamma_2 Y_{t-2} \dots + \gamma_i Y_{t-i} + \epsilon + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} \dots \theta_j \epsilon_{t-j}$$

A este modelo se le conoce como ARIMA(i,0,j).

Aunque el termino ϵ no tiene necesariamente una distribución normal, por el resto de documento se asumirá una distribución normal con media μ y varianza σ^2 , a menos que se especifique lo contrario.

Los procesos ARIMA resultan al considerar una serie de la forma:

$$Y_t = \alpha + Y_{t-1}$$

Tal que el proceso $Y_t - Y_{t-1}$ es un proceso ARMA.

1.2.2 Modelos Estado-Espacio

Tomado de [?] :

En un modelo estado espacio una serie de tiempo (multiple) observada y_1, \dots, y_t depende de un estado z_t , posiblemente no observado, que se comporta siguiendo un proceso estocástico. La relación entre y_t y z_t está dada por la *ecuación de medida*: [?]

$$y_t = H_t z_t + v_t$$

donde H_t es una matriz que puede o no depender del tiempo t y v_t es el error de observación, que se asume usualmente como un proceso de ruido. El vector de estado es generado como:

$$z_t = B_{t-1} z_{t-1} + w_{t-1}$$

La matriz B_t es una matriz de coeficientes que puede depender de t y w_t es un proceso de ruido. [?]

Es posible considerar los modelos estado-espacio una generalización de los modelo SARIMA y VARMA.

La población carcelaria en Colombia 1991 - 2017

2.1 Análisis exploratorio

El INPEC publica mensualmente la serie población carcelaria. Esta serie contiene la población carcelaria desde 1991, separada por situación jurídica (condenados, sindicados) y género.

La población carcelaria total entre 1991 y 2017 se ha cuadruplicado, al pasar de 32.036 a 128.125 internos. Ver figura 2.1. Aunque la mayoría de los internos son hombres, la población carcelaria femenina ha crecido a un ritmo aún más acelerado, al quintuplicar su población entre 1991 y 2017 (pasa de 1633 personas a 7800).

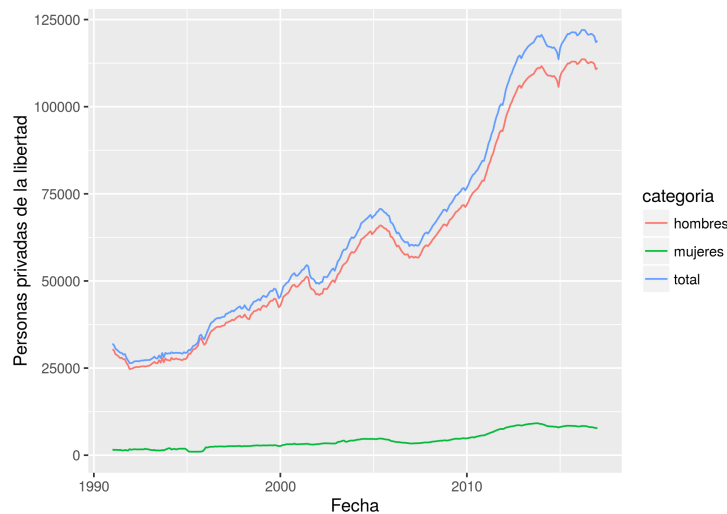


FIGURE 2.1. Población privada de la libertad 1991 - 2017

Fuente: INPEC

Elaboración propia

El incremento en la población carcelaria podría tomarse como un efecto del crecimiento de la población colombiana. Para validar este supuesto calculamos la tasa de encarcelamiento, que mide la cantidad de personas encarceladas por cada cien mil habitantes.

Este indicador pasó de 92 personas por cada cien mil habitantes en enero de 1991 a 242 en enero de 2016. Tal incremento se puede ver tanto en hombres como en mujeres. Ver figura 2.2.

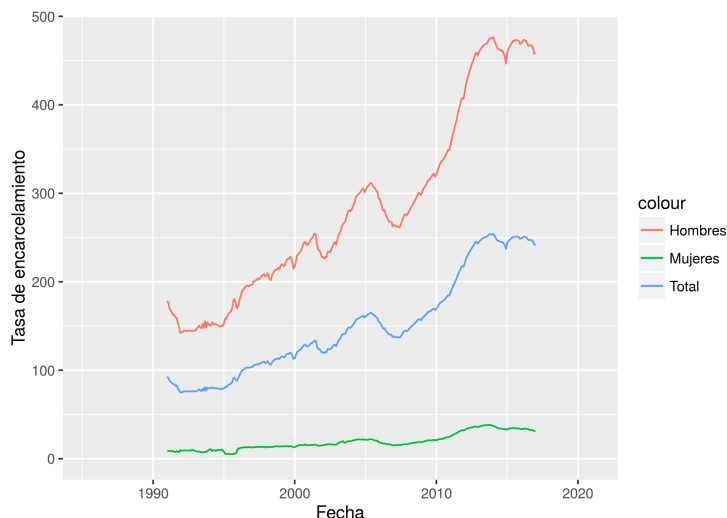


FIGURE 2.2. Tasa de encarcelamiento según género 1991 - 2017

Fuente: INPEC
Elaboración propia

La tasa de encarcelamiento es un indicador que varía según la edad y género, siendo más elevado en los hombres que en las mujeres y más algo en los hombres jóvenes, que en los hombres mayores (cita pendiente). Otra posible explicación al cambio en la tasa de encarcelamiento es un cambio en la pirámide poblacional en el periodo analizado. No obstante, no podemos confirmar o refutar esta hipótesis pues la serie de tiempo, contenida en los datos de libre acceso no se encuentra desagregada por edad.

2.2 El sistema penitenciario en Colombia

La población carcelaria se ve afectada por dos políticas, la política penitenciaria, que determina las condiciones de privación de la libertad y la política criminal que determina las causas de encarcelamiento y la duración de las penas. [5]

A la población privada de la libertad antes del juicio se le denomina población sindicada, y a aquellos que han sido juzgados y se encuentran cumpliendo la sentencia se les denomina población condenada. La evolución de la población según situación jurídica se puede observar en la figura 2.3

2.3 Identificación del modelo

Podemos modelar el sistema penitenciario de la siguiente manera:

$$\begin{aligned} S_t &= S_{t-1} + \alpha N_t - \gamma S_{t-1} \\ C_t &= C_{t-1} - \omega C_{t-1} + \beta \gamma S_{t-1} \end{aligned}$$

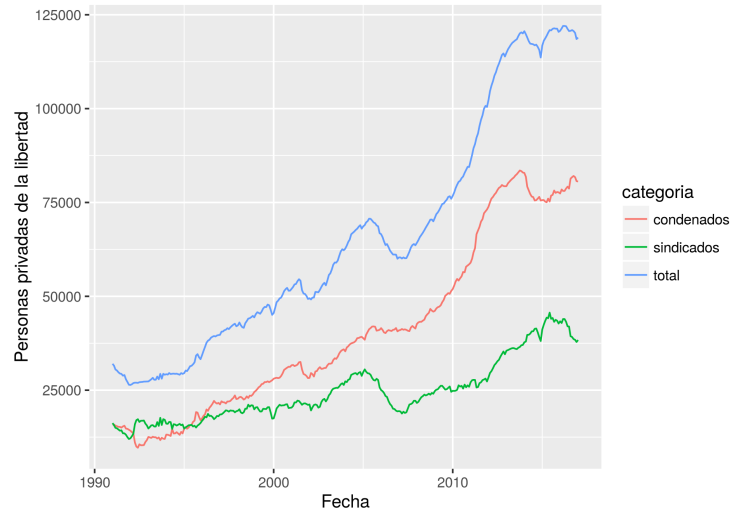


FIGURE 2.3. Población carcelaria por situación jurídica

Fuente: INPEC
Elaboración propia

N_t = población nacional en el periodo t
 S_t = población de sindicados en el periodo t
 C_t = población de condenados en el periodo t
 α = proporción de la población libre que ingresa al sistema carcelario
 γ = proporción de sindicados que es juzgada cada periodo
 β = proporción de sindicados que han sido encontrados culpables durante el juicio
 ω = proporción de condenados que cumplen su pena cada periodo.

Y esto es un problema interesante porque no tengo las series de tiempo de la transición!

CHAPTER 3

Modelos SARIMA

3.1 Marco teórico

3.2 Identificación del modelo

3.3 Proyecciones 2017 - 2020

3.4 Conclusiones

CHAPTER 4

Modelos estado-espacio

- 4.1 Marco teórico
- 4.2 Identificación del modelo
- 4.3 Simulación Monte-Carlo
- 4.4 Estimación de parámetros
- 4.5 Proyecciones 2017 - 2020
- 4.6 Conclusiones

Conclusiones

- La escritura de tesis utilizando L^AT_EX permite que se obtengan documentos de una presentación elegante, agradable, de una impresión incomparable, de escritura bastante simple en cuanto al texto técnico y de formulas matemáticas, junto con un manejo automático del formato de las partes de un documento y las referencias bibliográficas, desprendiéndose así de los detalles de edición que en otras herramientas, producen tantas frustraciones y dolores de cabeza.

Trabajo futuro

- Implementar y corregir todos aquellos errores que los usuarios de esta plantilla puedan encontrar, así como las sugerencias para la modificación de la plantilla que sean pertinentes.

Bibliography

- [1] Alfred Blumstein, Jacqueline Cohen, and Harold D. Miller, *Demographically disaggregated projections of prison populations*, Journal of Criminal Justice **8** (1980), no. 1, 1–26.
- [2] Peter J. Brockwell, *Time Series*, International Encyclopedia of Statistical Science, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 1601–1605.
- [3] Jacques J.F. Commandeur and Siem Jan Koopman, *An Introduction to State Space Time Series Analysis*, (2007), 189.
- [4] Departamento Administrativo Nacional de Estadística (DANE), *Comunicado de prensa*, 2009.
- [5] Departamento Nacional de Planeación, *Conpes 3828 POLÍTICA PENITENCIARIA Y CARCELARIA EN COLOMBIA*, (2015).
- [6] Institute for Criminal Policy Research, *Colombia / World Prison Brief*, 2016.
- [7] Instituto Nacional Penitenciario y Carcelario, *Plan de direccionamiento estratégico 2015-108, Misi{ó}n y Visi{ó}n*, 2016.
- [8] Ministry Justice, *Prison Population Projections 2014 - 2020, England and Wales*, Home Office Statistical Bulletin (2014), no. November, 31.
- [9] Ronald Demos Lee and Shripad Tuljapurkar, *Stochastic Population Forecasts for the United States: Beyond High, Medium, and Low*, Journal of the American Statistical Association **89** (1994), no. 428, 1175.
- [10] Todd D Minton, Bjs Statistician, Scott Ginder, Susan M Brumbaugh, Hope Smiley-Mcdonald, and Harley Rohloff, *Census of Jails: Population Changes, 1999 - 2013*, (2015).
- [11] David A Swanson and Jeff Tayman, *Subnational Population Estimates*, The Springer Series on Demographic Methods and Population Analysis, vol. 31, Springer Netherlands, Dordrecht, 2012.
- [12] Ruey S. Tsay, *Analysis of Financial Time Series*, John Wiley & Sons, Inc., Hoboken, NJ, USA, jan 2002.
- [13] Wai-Yin Wan¹, Steve Moffatt², Zachary Xie³, Simon Corben, and Don Weatherburn, *Forecastin prison populations using sentencing and arrest data*, Tech. report.