

Población Penitenciaria en Argentina 2002 a 2017

Tutor: Franco Camporeale

Nahuel Almeida, María Lucía Pappaterra, Javier ¿?

26 de mayo de 2019

1. Análisis estadístico de variables

Seleccionar un conjunto de al menos cuatro variables que resulten de interés.

1. Usar distintos tipos de gráficos para describir sus distribuciones
2. Analizar Outliers
3. Calcular estadísticos clásicos (media, mediana, moda, desviación estandar)

Las variables elegidas fueron: edad, duración de condena, diferencia entre fecha de condena y fecha de detención, cantidad de delitos por interno.

1.1. Edad

En la figura 1 mostramos la distribución de edades de los internos, junto con su suma acumulada. Podemos ver que la mayoría de los internos son personas jóvenes, y que el número de internos decae fuertemente con la edad. En el histograma se observa que hay registros de personas con edad cercana a 0 años. En realidad, se trata de 369 registros con edad igual a 0 años. Estos casos podrían tratarse de un error de carga, o podrían corresponder a niños nacidos dentro de los establecimientos. Observando en detalle esos registros, se puede ver que la segunda opción es improbable, dado que esos registros tienen información en otros campos que corresponden a internos que han cometido delitos. Los demás registros que podrían ser considerados como outliers son los correspondientes a menores de edad. En este caso, existen 8 registros de personas de 16 años y 11 registros de personas de 17 años. Como esta cantidad es baja y no afecta la estadística, decidimos incluirlos en el análisis.

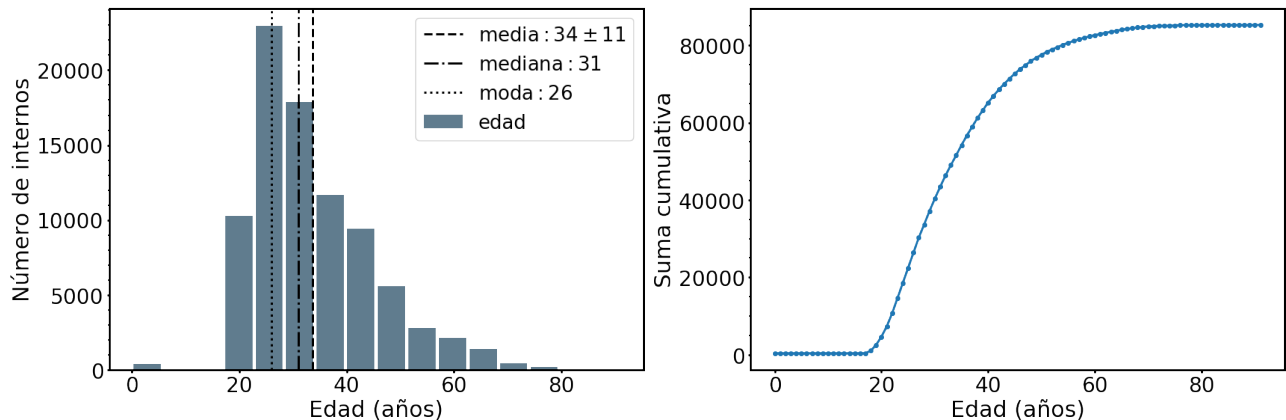


Figura 1: Histograma de edades de los internos, junto con su la suma acumulada.

1.2. Duración de condena

Para estudiar la duración de las condenas, es necesario utilizar dos campos: “duracion_condena_anios”, que indica la cantidad de años completos, y “duracion_condena_meses”, que indica la cantidad de meses restantes. Por ejemplo, una condena de 9 años y 5 meses figuraría en la base de datos como duracion_condena_anios: 9, duracion_condena_meses: 5. Para nuestro análisis, consideramos una nueva variable racional que sea la suma de los dos campos anteriores, medida en años.

El preprocesamiento fue el siguiente: Filtramos los registros con valor nulo en duracion_condena_anios (1) y con valor nulo en duracion_condena_meses (32). Luego, filtramos los registros con valor igual a 0 en ambos campos (3649).

En la figura 2 mostramos el histograma de duración de condenas, junto con su suma cumulativa complementaria. Podemos ver que la mayoría de las condenas son cortas (las tres medidas de centralidad calculadas son menores o iguales a 8 años), y que presentan un decaimiento aproximadamente exponencial a medida que se incrementa la duración.

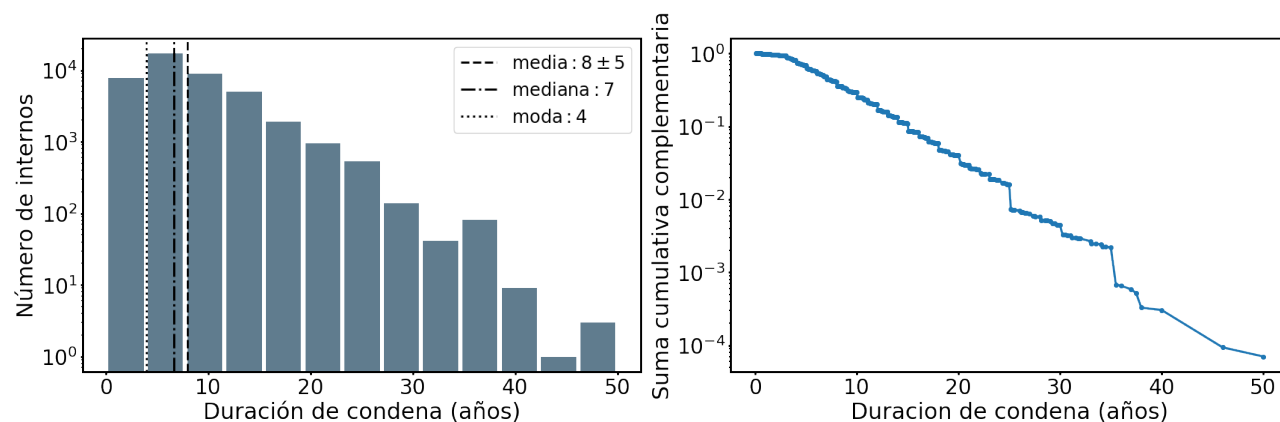


Figura 2: Histograma de la duración de condenas y su suma cumulativa complementaria.

1.3. Diferencia entre fecha de condena y fecha de detención

Una importante fracción de la población penal se encuentra detenida sin una condena firme. La situación más frecuente es la de la prisión preventiva. A continuación, mostramos un análisis sobre los tiempos transcurridos entre la detención y la aplicación de la condena.

Para hacer este análisis, utilizamos los campos “fecha_detencion” y “fecha_condenado”.

Preprocesamiento: Seleccionamos sólo los registros cuya situación penal sea “Condenado” (46405). Luego filtramos aquellos que tengan valores nulos para cualquiera de los dos campos de interés. En tercer lugar, filtramos los registros que tengan fechas posteriores a la fecha del censo (6409). Estos últimos casos son claramente outliers y corresponde descartarlos.

Al observar la distribución de diferencias de tiempo (figura 3) se puede ver que hay una gran cantidad de casos con diferencia de tiempo negativas. De hecho, la distribución está centrada en un valor cercano a cero, y decae aproximadamente tanto hacia valores positivos como negativos. Mientras que los valores positivos no parecen, a priori, extraños, es necesario interpretar los valores negativos. Para esto, hay que preguntarse si existe alguna condición en la que la fecha de condenado pueda preceder a la fecha de detención.

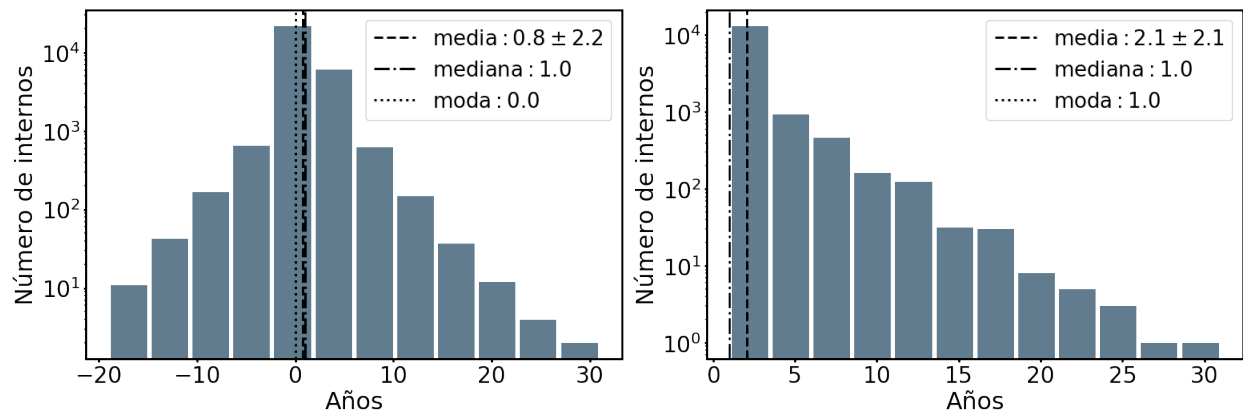


Figura 3: Histograma de los tiempos transcurridos entre la fecha de detención y la fecha de condena. A la izquierda, teniendo en cuenta diferencias de tiempo negativas, y a la derecha, considerando sólo las positivas.

Existen algunas hipótesis que podrían dar lugar a una situación de este tipo. La primera es que una persona podría haber recibido la condena sin estar detenida, o estando detenida en un establecimiento policial, cuyos datos no están dentro de la base de datos, y haber sido incorporada al servicio penitenciario un tiempo después. Podría suceder, incluso, que una persona se de a la fuga después de haber sido condenada, y antes de ser detenida. La segunda es que una persona puede tener múltiples causas. Se puede dar el caso en que una persona haya sido condenada por una causa y que, habiendo cumplido su condena, haya sido detenida por un segundo delito. En ese caso, podría ser que el registro muestre la fecha de condena de la primera causa, y la fecha de detención de la segunda. Para testear esta hipótesis, se podría estudiar si los registros para los cuales la diferencia de tiempos es negativa corresponde a registros de reincidentes. La tercera es que un condenado haya sido derivado de otro establecimiento y que figure como fecha de detención la fecha en la que ingresó al último establecimiento.

Para este primer análisis, vamos a considerar todos los valores negativos como outliers y descartarlos. Más adelante, buscaremos evaluar estas hipótesis. Al hacer esto, obtenemos la gráfica de la derecha en la figura 3.

1.4. Cantidad de delitos por interno

Cada interno puede haber cometido uno o más delitos. Cada delito está registrado en un campo distinto, y se registra hasta un máximo de 5 delitos, ordenados de mayor a menor gravedad¹. Aquí veremos cómo se distribuye el número de delitos por interno.

En el histograma de la figura 4 podemos ver el histograma con la cantidad de delitos por interno. Vemos que la gran mayoría (68915 registros) ha cometido un único delito, y que este número decae de manera aproximadamente exponencial para delitos mayores que 1.

¹No en todos los casos la clasificación está bien hecha. De hecho, existen registros donde el primer delito es nulo, y el segundo no lo es.

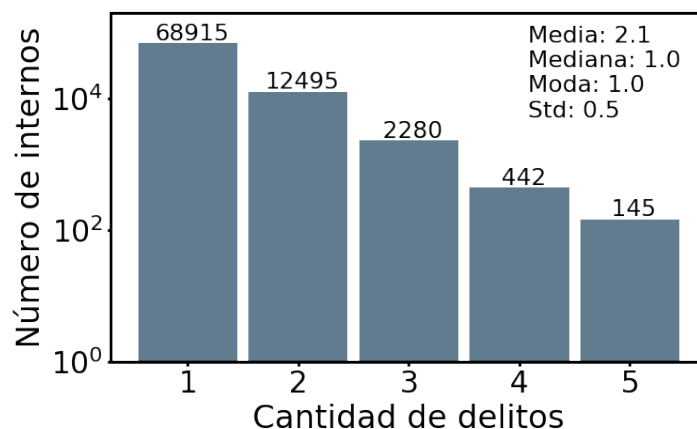


Figura 4: Histograma de la cantidad de delitos cometidos por cada interno.

2. Evolución de variables en el tiempo

Seleccionar dos variables y graficar cómo fueron cambiando desde 2002 a 2017. Para ello se tiene que utilizar el siguiente conjunto de datos: https://github.com/camporeale/Datos/raw/master/sneep_2002_2017_diplodatos.zip

Las variables elegidas fueron situación legal y participación en programas educativos.

2.1. Situación legal

Veamos cómo evoluciona la cantidad de internos condenados y procesados a lo largo de los años, y si nuestro análisis se condice con el de la nota de chequeado: <https://chequeado.com/ultimas-noticias/martin-casares-por-primera-vez-en-la-historia-en-2017-el-porcentaje-de-condenados-supero-al-de-procesados>

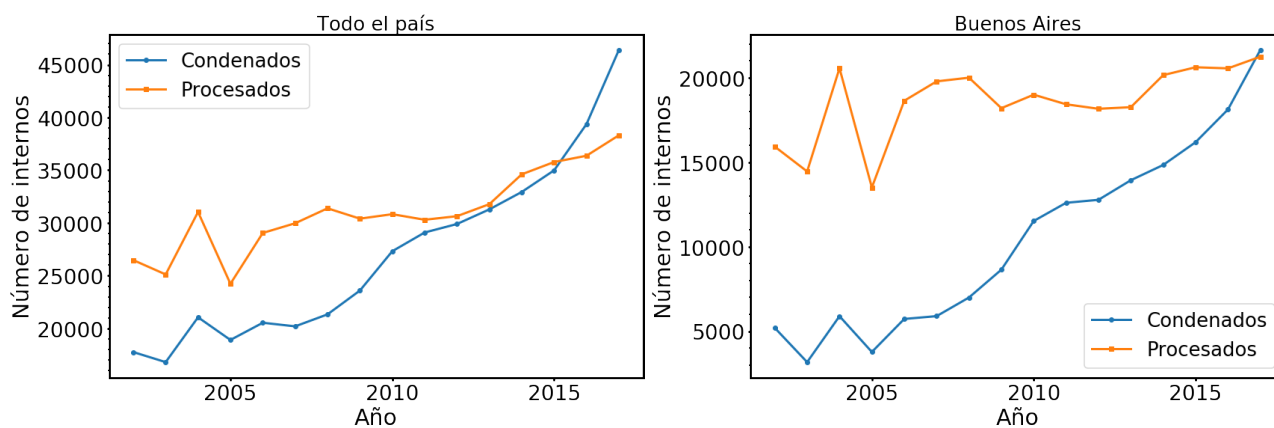


Figura 5

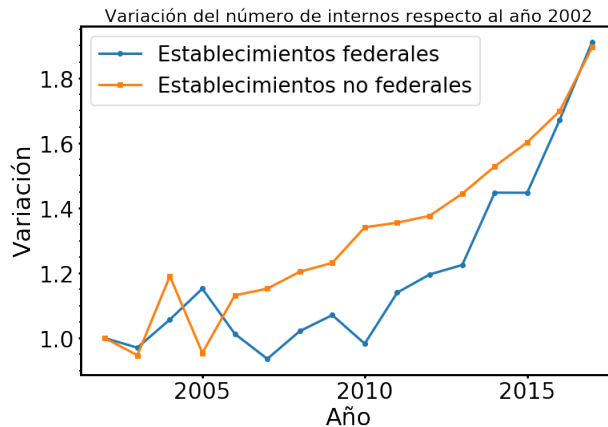


Figura 6

2.2. Participación en programas educativos

Queremos ver cómo varía la proporción de internos que participa en programas educativos a lo largo del tiempo.

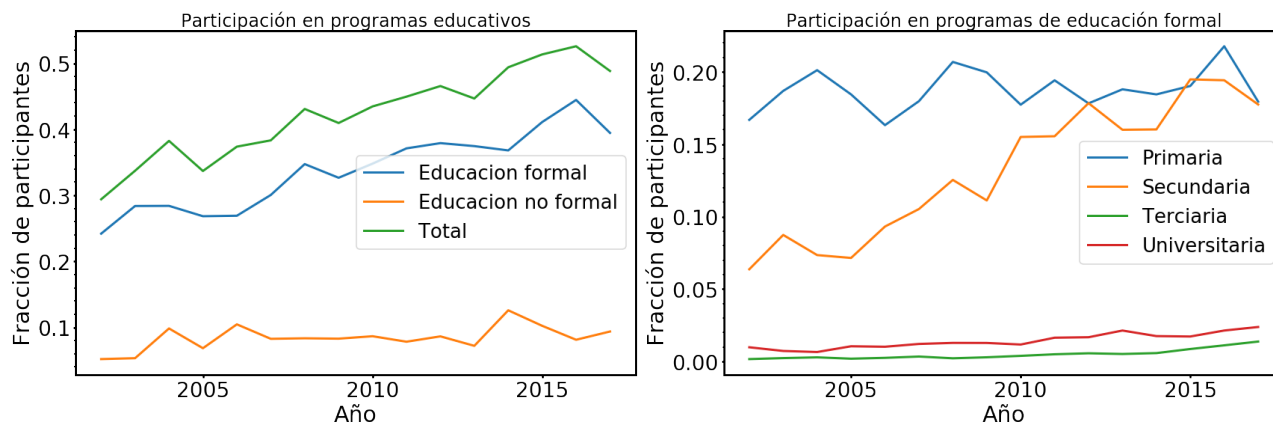


Figura 7

3. Análisis de probabilidades condicionales

Tomar al menos dos pares de variables y realizar un análisis del tipo:

- ¿Cuál es el porcentaje de internos argentinos y extranjeros? ¿Cómo varían esos números de acuerdo a si el establecimiento es federal o no federal?
- ¿Cuál es la probabilidad de que el interno haya sido lesionado en el último año dado que está en una prisión en Buenos Aires? ¿Y en Córdoba?
- ¿Cuál es la probabilidad de que se le otorguen salidas provisionarias dado que esté casado/a? ¿Y siendo soltero?

Porcentaje_internos		federales no_federales	
Argentina	94.192768	Argentina	78.600488 95.671692
Paraguay	1.990178	Boliviana	6.075400 0.657285
Peruana	1.131371	Brasileña	0.406835 0.078463
Boliviana	1.126671	Chilena	0.813670 0.375592
Chilena	0.413544	China	0.067806 0.029584
Uruguay	0.382998	Colombiana	0.664497 0.223811
Otras	0.278437	Ecuatoriana	0.176295 0.015435
Colombiana	0.261989	Española	0.081367 0.019294
Brasileña	0.106910	Inglesa	0.000000 0.001286
China	0.032896	Italiana	0.027122 0.023153
Ecuatoriana	0.029371	Nigeriana	0.013561 0.000000
Española	0.024672	Otras	1.125576 0.198086
Italiana	0.023497	Paraguay	6.305940 1.580829
Sudafricana	0.002350	Peruana	4.990507 0.765332
Inglesa	0.001175	Sudafricana	0.027122 0.000000
Nigeriana	0.001175	Uruguay	0.623813 0.360156

Figura 8