



Facultad de Estudios Superiores

Acatlán

Generaciones Simuladas con Técnicas Estadísticas y Variables Observadas

Acosta Rangel Hugo Rene
Barrera Hernández Daniel
De la Cruz Cruz Santiago
Guerrero Banderas David
Moranchel González José Manuel
Solis Alonso César Jonathan

Simulación Estocástica

Diciembre de 2017

GENERACIONES SIMULADAS CON TÉCNICAS ESTADÍSTICAS Y VARIABLES OBSERVADAS.

Simular de una población de nuevo ingreso a la carrera de actuaría, y después de 8 semestre analizar la población de los que se terminan todos los créditos.



Tabla de contenidos

1	Marco Teórico	5
1.1	Introducción	5
1.2	Objetivo	6
1.3	Justificación	6
1.4	Explicación del proyecto	6
1.5	Diagrama del proyecto	6
1.6	Bitácora de Actividades	7
1.7	Simulación de monte Carlo	7
2	Etapas de una simulación	9
3	Herramienta utilizada para la simulación	11
3.1	NOTEBOOK DOCUMENT	11
3.2	JUPYTER NOTEBOOK APP	11
3.3	KERNEL	11
3.4	NOTEBOOK DASHBOARD	12
4	Librerías	13
4.1	Statsmodels	13
4.2	Math	13
4.3	Numpy	13
4.4	Os	13
4.5	Pygal	14

4.6	Seaborn	14
4.7	Matplotlib	14
4.8	Fitter	14
4.9	Scipy	14
4.10	Warnings	14
4.11	Itertools	14
4.12	Pandas	14
5	Ajuste del modelo	15
6	Resultados del ajuste	21
6.1	Resultados	21
6.2	Análisis	25
7	Simulaciones	27
8	Resultados de Simulaciones	29
8.1	Aumentar numero de ingreso	29
8.2	Asesorías	29
9	Conclusión	33
	Artículos	35
	Sitios web	35



1. Marco Teórico

1.1 Introducción

El presente proyecto pretende analizar el índice de egresados y aprobación dentro de la licenciatura de Actuaría, haciendo el estudio y análisis de datos obtenidos directamente de la Facultad de Estudios Superiores Acatlán. Dado que este proyecto se centrará en la simulación a partir de datos ya conocidos, resulta fundamental entender el concepto de simulación, para ello citaremos a R. E. Shannon, el cual la define como : El proceso de diseñar un modelo de un sistema real y llevar a término experimentos con él, con la finalidad de comprender el comportamiento del sistema o evaluar nuevas estrategias -dentro de los límites impuestos por un cierto criterio o un conjunto de ellos - para el funcionamiento del sistema .

El interés por simular una generación nueva de alumnos nace del alto grado de reprobación observado a lo largo de las generaciones que han cursado la licenciatura y este proyecto no pretende únicamente recalcar esto, también tiene como objetivo prestar especial atención y llevar un seguimiento de aquellas materias en las que el resultado sea más alarmante. Tomando esto en cuenta, esta simulación se llevará a cabo con datos de las últimas generaciones desplegadas por profesor y por materia.

Esto resultará útil para analizar el aprovechamiento de los recursos destinados por la Universidad, ya que según datos de Cuenta Anual 2000-2016, Presupuesto 2017, UNAM cada alumno cuenta con un presupuesto de más de 50 mil pesos.

Así mismo, resultará útil para generaciones futuras que cursen la Licenciatura, ya que con esto podrán analizar las materias que son más propensas a ser reprobadas y prestar mayor atención en ellas, así como encontrar el camino óptimo para cursar la licenciatura, enfocados en el área de especialidad de su agrado.

1.2 Objetivo

Simular de una población de nuevo ingreso a la carrera de actuaría, y después de 8 semestre analizar la población de los que terminan todos los créditos.

1.3 Justificación

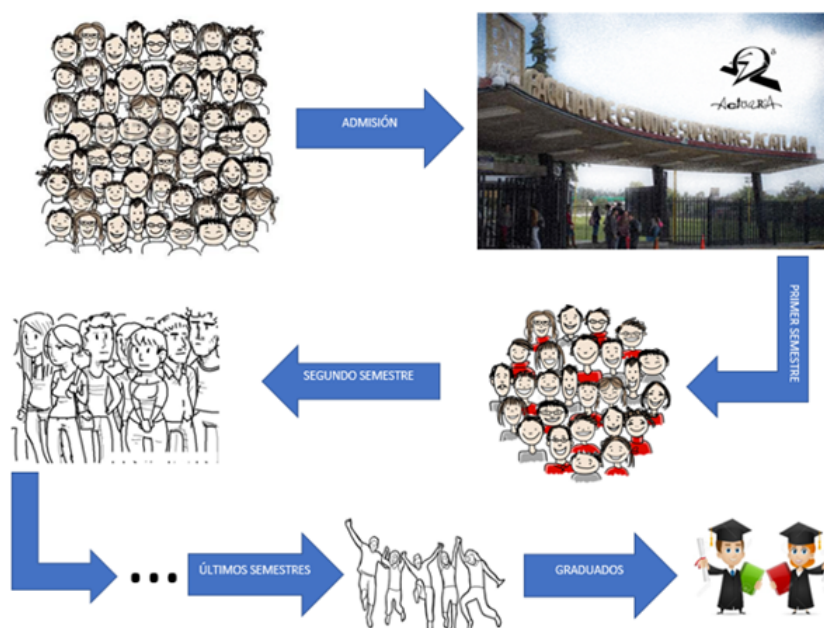
Actualmente la licenciatura es una carrera muy atractiva por lo que, una gran cantidad de personas está interesada en cursarla, pero una de las preguntas que se plantearían es: ¿Es difícil? ¿Qué tan difícil es? ¿Cuánto tiempo tendría que invertir en estudiarla? Con el fin de responder a esas inquietudes desarrollamos este proyecto.

1.4 Explicación del proyecto

Considerando la información proporcionada por la coordinación de actuaría, mediremos los índices de reprobación de cada materia y grupo.

Con esto empezaremos a simular una distribución de los grupos a quienes inscribiremos a 1er semestre, posteriormente simularemos cuántos pasaron por materia y con esos repetiremos el mismo procedimiento hasta terminar la carrera y nos fijaremos en la proporción de los que salen regulares respecto a los que entraron.

1.5 Diagrama del proyecto



1.6 Bitácora de Actividades



1.7 Simulación de monte Carlo

Nombrada así por la ciudad en Mónaco la cual es famosa por las apuestas, y el proceso de la simulación involucra generar variables aleatorias y describir comportamientos aleatorios.

La simulación es una poderosa herramienta estadística, usada principalmente para la difusión de neutrones en problemas como para la bomba atómica estudiado en el laboratorio científico Álamos en 1944. Monte Carlo se refiere usar números aleatorios como herramienta para calcular algo que no es aleatorio. Por ejemplo, sea X una variable aleatoria y denotamos por A , a su esperanza. Si nosotros podemos generar una muestra aleatoria de n número de veces X , entonces podemos crear la aproximación usando el estimador dado por la media.

Media Aritmética de una Muestra II
Datos NO Agrupados:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

\bar{X} : Media Aritmética
 X_i : i-ésimo valor observado
 n : Tamaño Muestra

calidadtotal@hotmail.com

9

La Ley de los grandes números nos dice que si hacemos tender n a infinito entonces nuestra estimación será nuestro valor esperado. Como X y la media son aleatorios y pueden ser diferentes cada vez que corramos nuestro programa, sin embargo, nuestro valor esperado no es aleatorio.

Hacemos énfasis en la distinción de Monte Carlo y Simulación, Simulación hace referencia a producir variables aleatorias con cierta distribución solo para observar su comportamiento. Por ejemplo, si queremos un modelo de un proceso aleatorio que produzca nubes, podemos generar un modelo que simplemente imágenes de nubes, más que de una cuestión de interés científica o computacional. En el momento en el que nos empezamos a preguntar por cuestiones cuantitativas, ya sea el tamaño promedio por nube o la probabilidad de que llueva, empezamos a movernos de Simulación a Monte Carlo, el motivo de hacer distinción es porque hay otras formas de definir A que hacen más fácil su estimación, conocido como reducción de varianza, ya que la mayor parte del error del estimador es estadístico. Reduciendo la varianza del estimador, reduces el error estadístico.

Frecuentemente decidimos entre Monte Carlo y métodos determinísticos, ya que usando un método de panel de integración para estimar el valor esperado de una variable aleatoria univariada con función de densidad $f(x)$, es más preciso debido a que el error de Monte Carlo es proporcional a $1/\sqrt{n}$ para n grande, lo cual da una aproximación precisa de $1/2$, es decir, que los métodos determinísticos son mejores que Monte Carlo en situaciones donde sean más prácticos.

Una característica favorable de Monte Carlo es que es posible estimar el orden de la magnitud del error estadístico, conocidos como barras de error debido a como se indican en graficas de resultados de Monte Carlo, las barras de error son esencialmente intervalos de confianza, otra característica favorable es que ideas ingeniosas pueden llevar a mejorar enormemente la eficiencia y precisión.



2. Etapas de una simulación

En el desarrollo de una simulación se pueden distinguir las siguientes etapas (Banks et al., 1996):

- **Formulación del problema.** En este paso debe quedar perfectamente establecido el objeto de la simulación. El cliente y el desarrollador deben acordar lo más detalladamente posible los siguientes factores: los resultados que se esperan del simulador, el plan de experimentación, el tiempo disponible, las variables de interés, el tipo de perturbaciones a estudiar, el tratamiento estadístico de los resultados, la complejidad de la interfaz del simulador, etc. Se debe establecer si el simulador será operado por el usuario o si el usuario sólo recibirá los resultados. Finalmente, se debe establecer si el usuario solicita un trabajo de simulación o un trabajo de optimización.
- **Definición del sistema.** El sistema a simular debe estar perfectamente definido. El cliente y el desarrollador deben acordar dónde estará la frontera del sistema a estudiar y las interacciones con el medioambiente que serán consideradas.
- **Formulación del modelo.** Esta etapa es un arte y será discutida más adelante. La misma comienza con el desarrollo de un modelo simple que captura los aspectos relevantes del sistema real. Los aspectos relevantes del sistema real dependen de la formulación del problema; para un ingeniero de seguridad los aspectos relevantes de un automóvil son diferentes de los aspectos considerados por un ingeniero mecánico para el mismo sistema. Este modelo simple se irá enriqueciendo como resultado de varias iteraciones.
- **Colección de datos.** La naturaleza y cantidad de datos necesarios están determinadas por la formulación del problema y del modelo. Los datos pueden ser provistos por registros históricos, experimentos de laboratorios o mediciones realizadas en el sistema real. Los mismos deberán ser procesados adecuadamente para darles el formato exigido por el modelo.
- **Implementación del modelo en la computadora.** El modelo es implementado utilizando algún lenguaje de computación. Existen lenguajes específicos de simulación que facilitan esta tarea; también, existen programas que ya cuentan con modelos implementados para casos especiales.
- **Verificación.** En esta etapa se comprueba que no se hayan cometido errores durante la

implementación del modelo. Para ello, se utilizan las herramientas de debugging provistas por el entorno de programación.

- **Validación.** En esta etapa se comprueba la exactitud del modelo desarrollado. Esto se lleva a cabo comparando las predicciones del modelo con: mediciones realizadas en el sistema real, datos históricos o datos de sistemas similares. Como resultado de esta etapa puede surgir la necesidad de modificar el modelo o recolectar datos adicionales.
- **Diseño de experimentos.** En esta etapa se decide las características de los experimentos a realizar: el tiempo de arranque, el tiempo de simulación y el número de simulaciones. No se debe incluir aquí la elaboración del conjunto de alternativas a probar para seleccionar la mejor, la elaboración de esta lista y su manejo es tarea de la optimización y no de la simulación. Debe quedar claro cuando se formula el problema si lo que el cliente desea es un estudio de simulación o de optimización. Teoría de Modelos y Simulación. Introducción a la Simulación.
- **Experimentación.** En esta etapa se realizan las simulaciones de acuerdo el diseño previo. Los resultados obtenidos son debidamente recolectados y procesados.
- **Interpretación.** Se analiza la sensibilidad del modelo con respecto a los parámetros que tienen asociados la mayor incertidumbre. Si es necesario, se deberán recolectar datos adicionales para refinar la estimación de los parámetros críticos.
- **Implementación.** Conviene acompañar al cliente en la etapa de implementación para evitar el mal manejo del simulador o el mal empleo de los resultados del mismo.
- **Documentación.** Incluye la elaboración de la documentación técnica y manuales de uso. La documentación técnica debe contar con una descripción detallada del modelo y de los datos; también, se debe incluir la evolución histórica de las distintas etapas del desarrollo. Esta documentación será de utilidad para el posterior perfeccionamiento del simulador.



3. Herramienta utilizada para la simulación

La herramienta utilizada para el proyecto es el programa Jupyter Notebook.

3.1 NOTEBOOK DOCUMENT

Los documentos del cuaderno (o *cuadernos*", en minúsculas) son documentos producidos por la aplicación Jupyter Notebook , que contiene tanto código de computadora (por ejemplo, python) como elementos de texto enriquecido (párrafos, ecuaciones, figuras, enlaces, etc.). Los documentos del cuaderno son documentos legibles por humanos que contienen la descripción del análisis y los resultados (figuras, tablas, etc.) así como documentos ejecutables que se pueden ejecutar para realizar análisis de datos.

3.2 JUPYTER NOTEBOOK APP

La aplicación Jupyter Notebook es una aplicación cliente-servidor que permite editar y ejecutar documentos portátiles a través de un navegador web. La aplicación Jupyter Notebook se puede ejecutar en un escritorio local que no requiere acceso a Internet (como se describe en este documento) o puede instalarse en un servidor remoto y acceder a través de Internet. Además de mostrar / editar / ejecutar documentos de cuaderno, la aplicación Jupyter Notebook tiene un "Tablero de instrumentos"(Panel de instrumentos portátil), un "panel de control"que muestra los archivos locales y permite abrir documentos de cuaderno o cerrar sus núcleos.

3.3 KERNEL

Un núcleo de portátil es un "motor computacional"que ejecuta el código contenido en un documento de Notebook . El núcleo IPython (IPython es un shell interactivo que añade funcionalidades extra al modo interactivo incluido con Python, como resaltado de líneas y errores mediante colores,

una sintaxis adicional para el shell, autocompletado mediante tabulador de variables, módulos y atributos; entre otras funcionalidades. Es un componente del paquete SciPy), al que se hace referencia en esta guía, ejecuta el código python.

Cuando abre un documento de Notebook, el kernel asociado se inicia automáticamente. Cuando se ejecuta el bloc de notas (celda por celda o con el menú Celda ->Ejecutar todo), el kernel realiza el cálculo y produce los resultados. Dependiendo del tipo de cálculos, el kernel puede consumir CPU y RAM significativas. Tenga en cuenta que la memoria RAM no se libera hasta que el kernel se apaga.

3.4 NOTEBOOK DASHBOARD

El tablero de instrumentos del portátil es el componente que se muestra primero cuando inicia la aplicación Jupyter Notebook. El tablero de instrumentos portátil se utiliza principalmente para abrir documentos portátiles, y para gestionar las ejecutan núcleos (visualizar y apagado). El panel de instrumentos del Notebook tiene otras características similares a un administrador de archivos, a saber, navegación de carpetas y cambio de nombre / eliminación de archivos.



4. Librerías

4.1 Statsmodels

Statsmodels es un modulo de python que provee clases y funciones para la estimación de diferentes modelos estadísticos, como también para pruebas estadísticas y exploración de datos estadísticos. Una extensa lista de resultados estadísticos disponibles para cada estimador. Los resultados son puestos a prueba contra paqueterías existentes para asegurar de que están bien.

4.2 Math

Este modulo siempre esta disponible, el cual proporciona acceso a las funciones matemáticas definidas por C standard. Estas funciones no pueden ser usadas con números complejos, para lo cual hay que usar cmath si se requieren funciones para números complejos.

4.3 Numpy

Es una paquetería fundamental para computación científica en python. Proporciona colecciones multidimensionales de objetos, varios objetos derivados como matrices, y operaciones sobre colecciones incluyendo manipulación matemática lógica y de forma como ordenar y seleccionar, también transformadas de Fourier discretas, álgebra lineal básica, operaciones estadísticas básicas, simulación aleatoria, entre otras.

4.4 Os

Este modulo proporciona funciones para leer, abrir y crear documentos, en general para manipular y diseñar documentos a través de Python.

4.5 Pygal

Esta paquetería proporciona una forma mas visual de representar los datos a treves de graficas de barras, de caja, de pie, histogramas, pirámides entre otros.

4.6 Seaborn

Seaborn es una librería para hacer mas atractivos e informativos los graficos estadísticos en python.

4.7 Matplotlib

Matplotlib es una librería par ahacer graficos en 2d de colecciones de python. Esta diseñado para crear simples graficos con poco codigos.

4.8 Fitter

Es una paquetería que proporciona una simple clase para descubrir de que distribución proviene tus datos. Utiliza la paquetería scipy para probar 80 distribuciones y permite graficar para verificar los resultados.

4.9 Scipy

Es una colección de algoritmos matemáticos, funciones creadas en la extensión de Numpy. Proporciona al usuario comandos y clases para manipular y visualizar datos.

4.10 Warnings

Util para manipular los mensajes de advertencia al usuario.

4.11 Itertools

El modulo estandariza un conjunto de herramientas raoidas y de eficiencia en memoria que son utiles por ellas mismas o en conjunto. Juntas forman una algebra iterativa haciendo posible construir herramientas especializadas eficientemente en Python puro.

4.12 Pandas

Proporciona herramientas para estructura de datos y análisis de datos en Python.

5. Ajuste del modelo

Como proceso y análisis en nuestro proyecto realizamos lo siguiente para el caso de los ingresos:

- Graficamos el comportamiento de los ingresos y vemos que toma una tendencia. Ver Figura 5.1
- Teníamos que simular los Ingresos por lo que usamos fitter para poder ajustar una distribución a los datos. Notamos que el ajuste no era el deseado ya que distaba mucho de los datos reales por lo que optamos por usar otro método para ajustarlos. Ver Figura 5.2
- Optamos por usar el método Arima, se trata de un modelo de Series de tiempo que en base a los datos pasados y con el uso de variaciones y regresiones podemos hacer predicciones a futuro. Conforme tenia mas datos pasados la predicción se acercaba mas a los datos reales. Ver Figura 5.3
- Graficamos la tendencia central de los datos. Primero, Calculamos los coeficientes del ajuste ($Ax+B$).Luego, Calculamos los coeficientes de correlación. Después, Dibujamos los datos para poder visualizarlos y ver si sería lógico considerar el ajuste usando un modelo lineal. Por

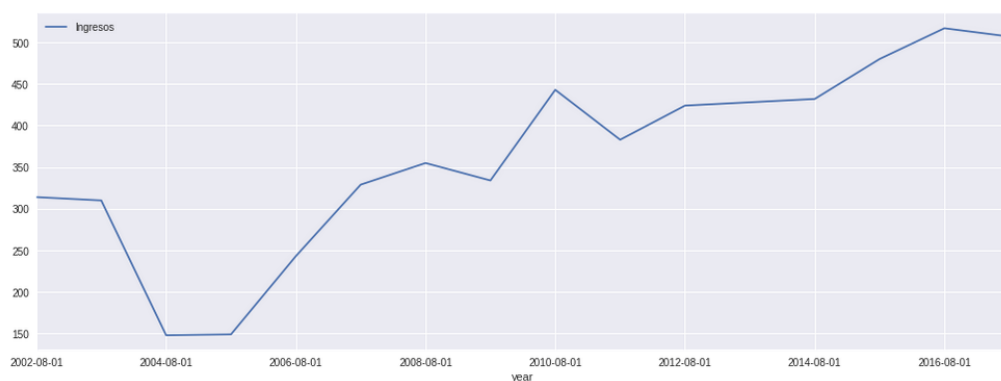


Figura 5.1: Histórico


```

                sumsquare_error
johnsonsb      0.003873
weibull_min    0.004156
frechet_r      0.004156
exponpow       0.004228
lognorm        0.004229
CPU times: user 580 ms, sys: 0 ns, total: 580 ms
Wall time: 574 ms

```

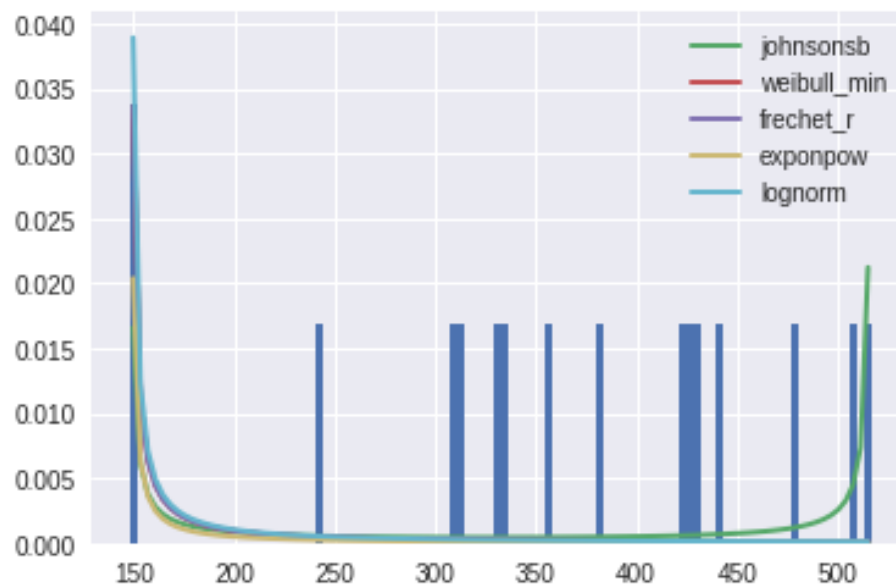


Figura 5.2: Ajuste de distribuciones

```

>Predicted=336.857, Expected=443
>Predicted=459.125, Expected=383
>Predicted=390.667, Expected=424
>Predicted=435.000, Expected=428
>Predicted=438.364, Expected=432
>Predicted=441.833, Expected=480
>Predicted=492.769, Expected=517
>Predicted=531.500, Expected=507
RMSE: 51.120

```

Figura 5.3: Errores

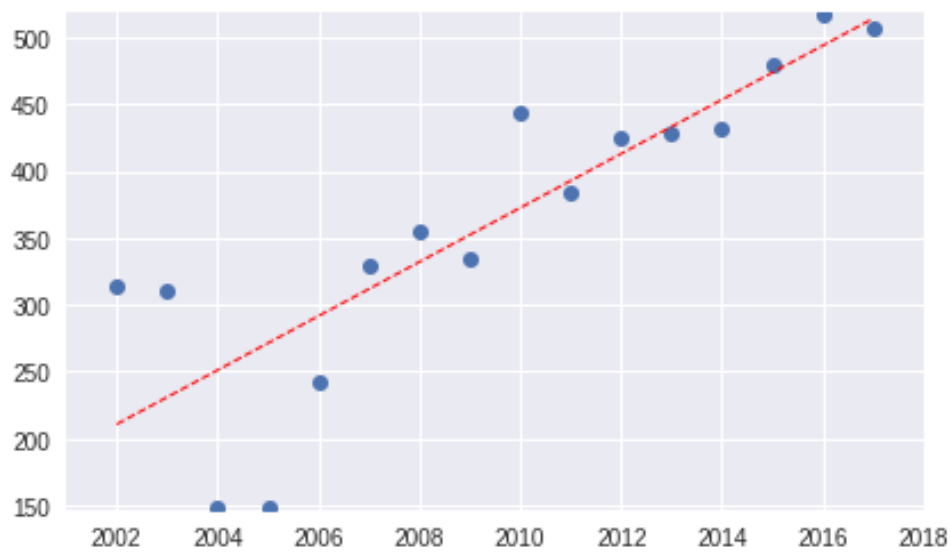


Figura 5.4: Tendencia central

último, usamos una distribución Normal para ver donde quedaría cada punto que intentamos predecir. Ver Figura 5.4

Para el caso de los egresos:

- Graficamos el comportamiento de los egresados y los comparamos con los titulados. Ver Figura 5.5
- De manera parecida a como simulamos los ingresos, ahora predecimos los egresos. Ver Figura 5.6
- La predicción fue bastante satisfactoria ya que en ocasiones llega a ser exacta, notamos de igual manera que entre mas datos tenemos menos varían con el dato real.
- Graficamos la tendencia central de los datos de egresados. Ver Figura 5.7
- Ejemplo de simulación de materias de primer semestre. Aprobado(1), no aprobado(0). Ver Figura 5.8

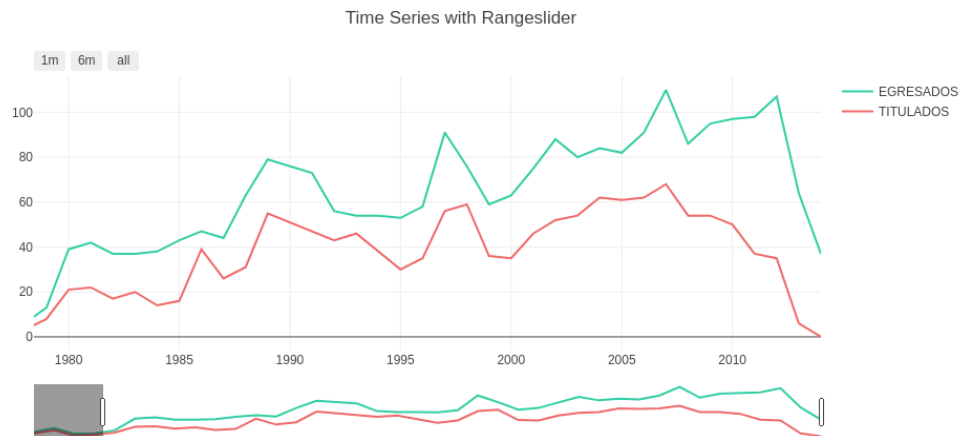


Figura 5.5: Histórico

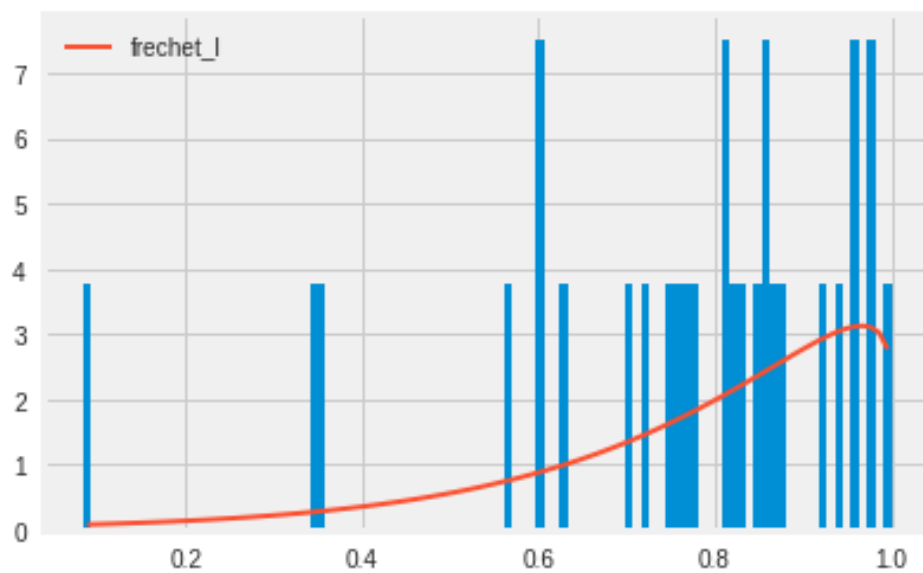


Figura 5.6: Ajuste de distribuciones

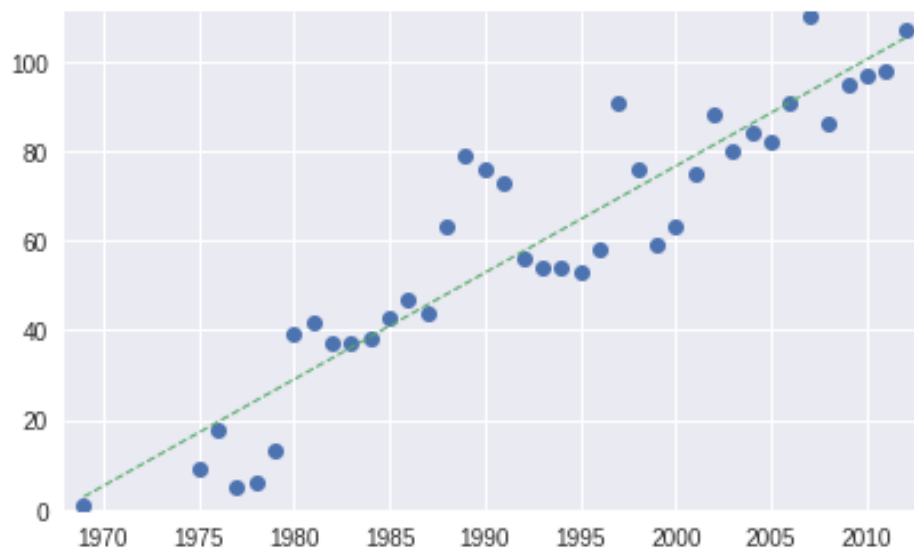


Figura 5.7: Tendencia Central

id_alumno	grupo	CALCULO DIFERENCIAL E INTEGRAL I	ALGEBRA SUPERIOR I	GEOMETRIA ANALITICA I	ALGORITMOS Y PROGRAMACION	SEGURO DE VIDA
0	1	6	0	0	1	1
1	2	3	0	0	1	1
2	3	2	0	0	1	1
3	4	7	1	0	1	1
4	5	2	1	0	1	1
5	6	4	0	0	1	0
6	7	4	1	1	1	0
7	8	2	0	0	1	1
8	9	5	0	1	0	1
9	10	5	0	1	1	1
10	11	3	0	1	0	0
11	12	7	0	0	0	1
12	13	5	0	1	0	1
13	14	5	0	0	1	1
14	15	2	0	0	0	1
15	16	1	1	0	1	1
16	17	2	1	1	0	1
17	18	4	1	0	1	0
18	19	2	1	0	1	1
19	20	5	0	1	0	1
20	21	7	0	0	0	1
21	22	6	0	0	0	1
22	23	1	0	1	0	1
23	24	7	1	0	1	1
24	25	4	0	0	1	0
25	26	6	1	0	0	1

Figura 5.8: Ejemplo Simulación



6. Resultados del ajuste

6.1 Resultados

Los resultados obtenidos al final de la simulación tuvieron como base principal el plan de estudios vigente en la licenciatura de Actuaría en la Facultad de Estudios Superiores Acatlán. El cual se puede visualizar en la Figura 6.1

Partiendo de los grupos asignados al principio de la generación y el número de grupos a lo largo de los semestres, considerando que van reduciendo el tamaño de alumnos regulares y por lo tanto el número de grupos abiertos, los alumnos son repartidos de forma uniforme a cada uno de los grupos, así mismo cada grupo se le asigna una probabilidad por materia como se muestra en la Figura 6.2

Teniendo la probabilidad por grupo, se determina si los alumnos que fueron asignados a ese grupo tuvieron una resultado aprobatorio esto con las probabilidades mencionadas anteriormente, de tal forma que se crea un historial académico por alumno en cada semestre como se puede ver en la Figura 6.3

Teniendo como filtro la seriación determinada en el plan de estudios, los alumnos se vuelven irregulares a lo largo de los semestres.

Después de haber simulado la misma generación n veces, los resultados que nos arroja de acuerdo al filtro mencionado anteriormente:

- Si se hacen 10 simulaciones la media queda con valor de 18 y nos arroja un histograma como se puede ver en la Figura 6.4
- Si se hacen 50 simulaciones la media queda con valor de 20 y nos arroja un histograma como se puede ver en la Figura 6.4
- Si se hacen 100 simulaciones la media queda con valor de 19 y nos arroja un histograma como se puede ver en la Figura 6.4
- Si se hacen 1000 simulaciones la media queda con valor de 19 y nos arroja un histograma como se puede ver en la Figura 6.4

MATERIA	SEMESTRE	MATERIA	SEMESTRE
CALCULO DIFERENCIAL E INTEGRAL I	1	CALCULO DIFERENCIAL E INTEGRAL II	2
ALGEBRA SUPERIOR I	1	ALGEBRA SUPERIOR II	2
GEOMETRIA ANALITICA I	1	GEOMETRIA ANALITICA II	2
ALGORITMOS Y PROGRAMACION	1	PROBLEMAS SOCIOECON. MEX. ACTUAL	2
SEGURO DE VIDA	1	BASES DE DATOS	2
		SEGURO DE DAÑOS	2
		INGLES I	2
(a) Primer semestre		(b) Segundo semestre	
MATERIA	SEMESTRE	MATERIA	SEMESTRE
CALCULO DIFERENCIAL E INTEGRAL III	3	CALCULO DIFERENCIAL E INTEGRAL IV	4
ALGEBRA LINEAL I	3	ALGEBRA LINEAL II	4
CONTABILIDAD	3	INVESTIGACION DE OPERACIONES I	4
PROBABILIDAD I	3	ESTADISTICA I	4
MATEMATICAS FINANCIERAS I	3	MATEMATICAS FINANCIERAS II	4
INGLES II	3	INGLES III	4
(c) Tercer semestre		(d) Cuarto semestre	
MATERIA	SEMESTRE	MATERIA	SEMESTRE
ECUACIONES DIFERENCIALES	5	ANALISIS MATEMATICO I	6
PROBABILIDAD II	5	ECONOMIA MATEMATICA I	6
ESTADISTICA II	5	PROCESOS ESTOCASTICOS I	6
FINANZAS CORPORATIVAS	5	ESTADISTICA III	6
MATEMATICAS ACTUARIALES I	5	APLICACION A LAS MATEMATIC.FINAN	6
INGLES IV	5	MATEMATICAS ACTUARIALES II	6
		INGLES V	6
(e) Quinto semestre		(f) Sexto semestre	
MATERIA	SEMESTRE	MATERIA	SEMESTRE
ANALISIS NUMERICO	7	SEMINARIO DE INVESTIGACION	8
DEMOGRAFIA MATEMATICA I	7	ADMINISTRACION DE RIESGOS	8
TEORIA DEL RIESGO I	7	MATEMATICAS ACTUARIALES PENSIONE	8
ADMINISTRACION ACTUARIAL	7	OPTATIVA_3	8
INGLES VI	7	OPTATIVAS_4	8
OPTATIVA_1	7	OPTATIVAS_5	8
OPTATIVA_2	7		
(g) Septimo semestre		(h) Octavo semestre	

Figura 6.1: Plan de estudios

MATERIA	proba_aproba
ECUACIONES DIFERENCIALES	0.230769
PROBABILIDAD II	0.756098
ESTADISTICA II	0.855072
FINANZAS CORPORATIVAS	0.911765
MATEMATICAS ACTUARIALES I	0.666667
INGLES IV	0.851852
MATERIA	proba_aproba
ECUACIONES DIFERENCIALES	0.500000
PROBABILIDAD II	1.000000
ESTADISTICA II	0.442857
FINANZAS CORPORATIVAS	0.942857
MATEMATICAS ACTUARIALES I	0.339286
INGLES IV	0.571429
MATERIA	proba_aproba
ECUACIONES DIFERENCIALES	0.410959
PROBABILIDAD II	0.843750
ESTADISTICA II	1.000000
FINANZAS CORPORATIVAS	0.937500
MATEMATICAS ACTUARIALES I	0.418182
INGLES IV	0.869565
MATERIA	proba_aproba
ECUACIONES DIFERENCIALES	0.431373
PROBABILIDAD II	0.833333
ESTADISTICA II	0.507246
FINANZAS CORPORATIVAS	1.000000
MATEMATICAS ACTUARIALES I	0.371429
INGLES IV	0.571429
MATERIA	proba_aproba
ECUACIONES DIFERENCIALES	0.480000
PROBABILIDAD II	0.500000
ESTADISTICA II	0.560000
FINANZAS CORPORATIVAS	0.851852
MATEMATICAS ACTUARIALES I	0.706897
INGLES IV	0.851852
MATERIA	proba_aproba
ECUACIONES DIFERENCIALES	0.478261
PROBABILIDAD II	0.380952
ESTADISTICA II	0.400000
FINANZAS CORPORATIVAS	0.942857
MATEMATICAS ACTUARIALES I	0.706897
INGLES IV	0.806452
MATERIA	proba_aproba
ECUACIONES DIFERENCIALES	0.618182
PROBABILIDAD II	0.819672
ESTADISTICA II	0.589041
FINANZAS CORPORATIVAS	1.000000
MATEMATICAS ACTUARIALES I	0.777778
INGLES IV	0.571429

Figura 6.2: Probabilidades de aprobación por grupos

id_alumno	grupo		ANALISIS MATEMATICO I	ECONOMIA MATEMATICA I	PROCESOS ESTOCASTICOS I	ESTADISTICA III	APLICACION A LAS MATEMATIC.FINAN	MATEMATICAS ACTUARIALES II	INGLES V
0	1	5	0	1	0	1	0	1	1
1	2	4	1	1	0	1	1	1	0
2	3	5	1	1	0	1	1	0	1
3	4	5	1	1	0	1	0	1	1
4	5	1	1	1	1	1	1	1	1
5	6	2	0	0	1	1	1	0	1
6	7	7	1	0	1	1	0	1	1
7	8	3	0	0	1	1	1	0	0
8	9	1	1	1	1	1	1	1	1
9	10	6	0	0	1	0	1	1	1

Figura 6.3: Historial alumno

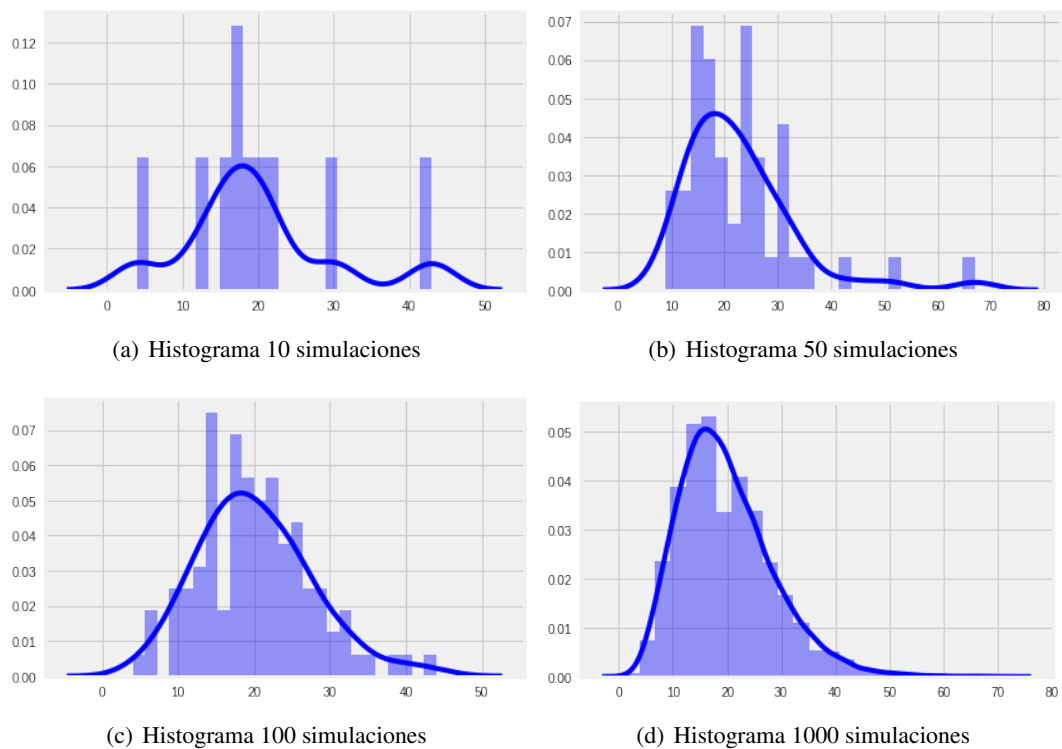


Figura 6.4: HISTOGRAMAS

6.2 Análisis

Analizando los resultados obtenidos, podemos concluir que cursar la licenciatura en tiempo y forma no es trabajo sencillo ya que, de una generación inicial de 500 alumnos, con una desviación de ± 30 , al final de 4 años, solo 30 (± 5 alumnos) logran aprobar las materias y terminar como es debido.

Así mismo, uno de los grandes obstáculos detectados en el curso “natural” de una generación nueva, fueron los filtros utilizados (que representan la seriación obligatoria de las materias); estos filtros ajustan probabilidades al evento de aprobar o no una materia y así, causaron enormes bajas del alumnado a lo largo de la licenciatura y hacen que, desde el primer semestre, la generación “regular” se reduzca a menos de la mitad. Por lo que si se pretende continuar y terminar la carrera con regularidad, es esencial concluir el primer semestre sin adeudar materias.

Otra cosa que se pudo observar con la simulación fue que, al igual que en el sistema real, las materias con una seriación más larga se les adjudicaban mayor índice de reprobación, por lo que se puede concluir que esta tendencia seguirá generación tras generación, por lo que los coordinadores de la carrera como el alumnado debe de poner mayor atención a estas materias y buscar alguna solución a esto si se desea que esta tendencia cambie.

Por otro lado, la herramienta de la simulación es de gran ayuda para poder responder cuestiones en un principio desconocidas, a través de datos registrados con anterioridad y aunado a esto viene el acceso a estos datos históricos ya que, mientras contemos con más datos, la simulación será más acertada y cercana a situaciones más “realistas”.

Por supuesto no podemos dejar de lado el uso de otras herramientas útiles para el desarrollo de este proyecto, como lo son el Machine Learning, y el Forecast a través de modelos ARIMA. Estas y otras herramientas, respaldadas por ciencias como la estadística y la probabilidad, nos permiten analizar y concluir, siempre con cierto nivel de confianza, acerca de situaciones parcialmente desconocidas.



7. Simulaciones

Después de haber analizado los resultados obtenidos del ajuste y siguiendo con nuestro objetivo, propondremos modelos en los cuales trataremos de provocar un aumento en el número de egresados, el cual tendrá una relación notable respecto al número de titulados. Los casos que tenemos contemplados para este estudio son:

- Aumentar número de ingresos: Como caso básico aumentaremos la matrícula de ingreso en la Licenciatura en Actuarial para ver que tipo de resultados tenemos.
- Asesorías: Tomando en cuenta un análisis realizado a la materia de Cálculo Diferencial e Integral I, ya que en este semestre hace unos años se implementó un sistema llamado PPMAIR, en el cual se le otorga un tipo de asesorías a los grupos de nuevo ingreso, esto con ayuda de estudiantes que se encuentren en semestres finales, se encontró que el aumento de aprobación de esta materia es de aproximadamente 5 %, si consideramos este aumento de aprobación en todas las materias de la carrera podría ser algo significativo. Esto repercutiría sobre todo en la seriación más extensa y con mayor problemática de la carrera, la cual es:
 - Cálculo Diferencial e Integral I
 - Cálculo Diferencial e Integral II
 - Probabilidad I
 - Estadística I
 - Probabilidad II

Y para el caso de :

- Procesos Estocásticos I
- Teoría del Riesgo I
- Administración de Riesgos

Se podría pedir apoyo a ex alumnos egresados de esta Licenciatura.



8. Resultados de Simulaciones

8.1 Aumentar numero de ingreso

Teniendo una media de egresados inicial de 19,

- En promedio se espera que para la generación de 2015 con aumento de 0, salgan: 20
- En promedio se espera que para la generación de 2015 con aumento de 50, salgan: 22
- En promedio se espera que para la generación de 2015 con aumento de 100, salgan: 25
- En promedio se espera que para la generación de 2015 con aumento de 150, salgan: 27
- En promedio se espera que para la generación de 2015 con aumento de 200, salgan: 29
- En promedio se espera que para la generación de 2015 con aumento de 250, salgan: 32

Ver Figura 8.1

En este estudio al obtener los resultados una vez variando el número de ingresos, y con ayuda de una regresión lineal obtuvimos que los resultados obtenidos se comportan aproximadamente de la siguiente manera:

$$y = 0,5X + 20$$

Lo cual se comporta de la forma (Ver Figura 8.2)

8.2 Asesorías

Teniendo una media de egresados inicial de 19,

- En promedio se espera que para la generación de 2015 con aumento de .00, salgan: 19
- En promedio se espera que para la generación de 2015 con aumento de .05, salgan: 31
- En promedio se espera que para la generación de 2015 con aumento de .10, salgan: 45
- En promedio se espera que para la generación de 2015 con aumento de .15, salgan: 74
- En promedio se espera que para la generación de 2015 con aumento de .20, salgan: 100

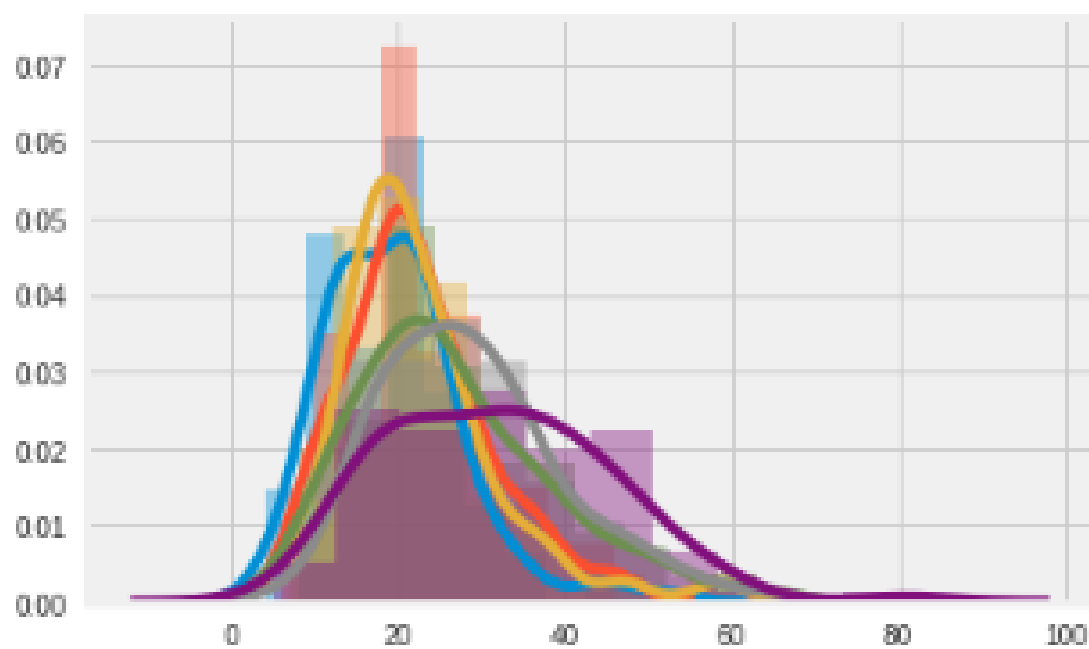


Figura 8.1: Aumento de matrícula

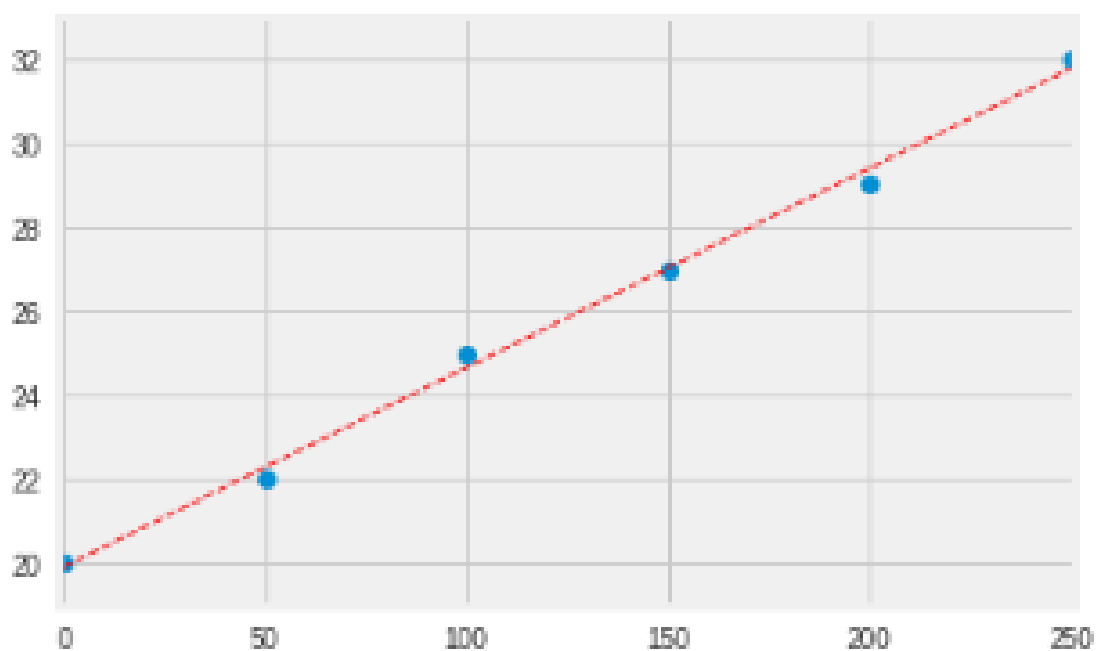


Figura 8.2: Comportamiento de egresados de acuerdo a el aumento por asesorías

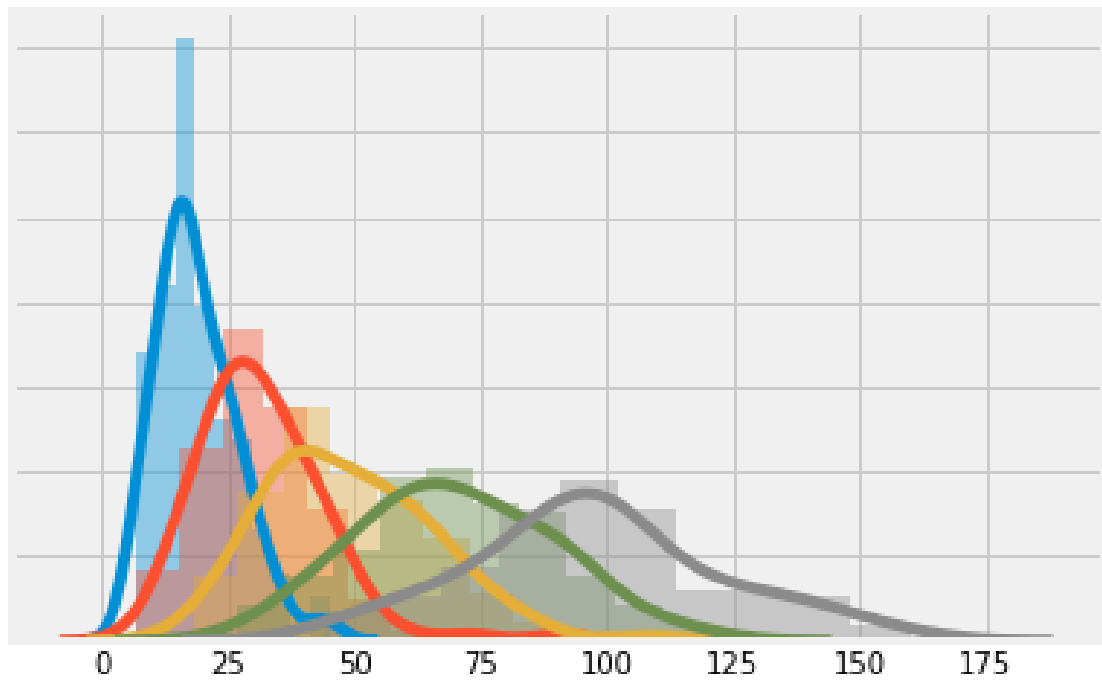


Figura 8.3: Cambio gráfico al aumentar 5 % de aprobación

Ver Figura 8.3

En este estudio al obtener los resultados una vez variando las probabilidades de aprobar, y con ayuda de interpolación de Lagrange obtuvimos que los resultados obtenidos se comportan aproximadamente de la siguiente manera:

$$20000,00000000005X^4 - 6000,00000000012X^3 + 1350,00000000002X^2 + 205,0X + 19,0$$

Lo cual se comporta de la forma (Ver Figura 8.4)

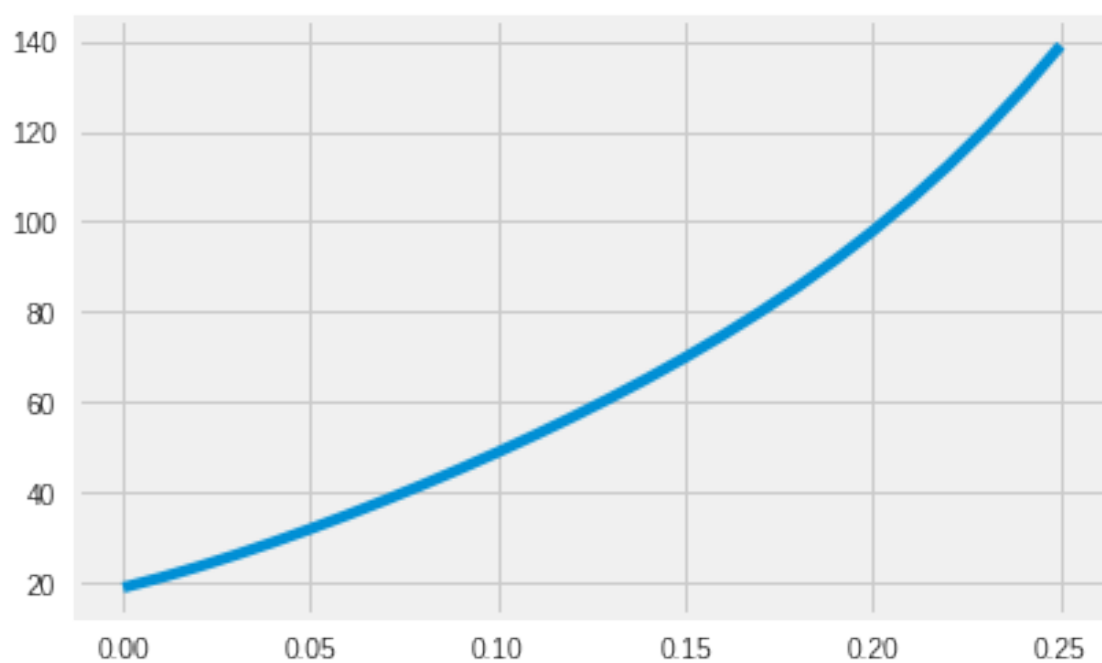


Figura 8.4: Comportamiento de egresados de acuerdo a el aumento por asesorías



9. Conclusión

Al obtener los resultados de nuestras simulaciones se puede observar que en el caso de aumentar el índice de aprobación, el número de egresados va creciendo en un número considerable, de hecho, aunque se eleve en un mismo porcentaje el índice de aprobación (en este caso de 5 % en 5 %) el número de egresados que se aumento no es el mismo, si no que fue aún mayor que en el caso anterior y esta tendencia sigue cada vez que se aumenta dicho índice.

En la simulación de aumentar el número de alumnos que ingresan a la carrera de Actuaría, nos encontramos que al aumentar inicialmente en aproximadamente un 20 % dicho número de ingresos, tendremos un aumento de egresados muy pobre, pues solo egresaron 4 alumnos más de los que habrían egresado sin aumentar el número de ingresos a la carrera, y esta tendencia sigue aún cuando se aumenta más la matrícula de ingresos.

Tomando en consideración los casos anteriores de simulación antes mencionados, si lo que se pretende es tener un mayor número de egresados año con año, es mejor invertir en mejorar el índice de aprobación de la carrera de Actuaría, a aumentar el número de alumnos que estudian esta, ya que hay una diferencia bastante considerable de un caso con otro respecto al aumento de egresados, pues al tener un incremento en el índice de aprobación se obtienen hasta el triple de resultados favorables con respecto al otro caso, y la distancia entre estos va aumentando cada vez que se hace un aumento en el índice de aprobación y número de egresados. Incluso, teniendo como objetivo tener más egresados de esta carrera, económicamente para la universidad sería bastante más favorable gastar cierto presupuesto en asesorías o talleres para mejorar el índice de aprobación que gastar en el mantenimiento de un aumento considerable en la matrícula de la carrera.



Bibliografía

Artículos

Series Temporales: Modelo ARIMA, Universidad Autónoma de Madrid, Santiago de la Fuente. El Financiero

Lagrange Interpolation, Jim Lambers, Fall Semester 2010-11, Lecture 5 Notes.

Método de Montecarlo, Área de Estadística e Investigación Operativa, Licesio J. Rodríguez-Aragón.

Sitios web

Simulación Montecarlo, Consultado el 23/11/2017, Obtenido de: https://previa.uclm.es/profesorado/licesio/Docencia/mcoi/Tema4_guion.pdf

Módulos y librerías en python, Consultado el 23/11/2017, Obtenido de: <http://docs.python.org.ar/tutorial/3/modules.html>

Modelos y tipos de simulación, Consultado el 25/11/2017, Obtenido de: https://prezi.com/_1vawverb61m/tipos-y-modelos-de-simulacion/

Información de jupyter, Consultado el 25/11/2017, Obtenido de: <http://jupyter-notebook-beginner-guide.readthedocs.io/>

Ayuda en elementos de programación, Consultado el 20/11/2017, Obtenido de: <https://stackoverflow.com/>

Herramienta online LaTeX, Obtenido de: <https://www.overleaf.com/>