

**DISEÑO DE EXPERIMENTOS APLICADO AL ANÁLISIS DE LA DINÁMICA DEL
FRAUDE ACADÉMICO EN LOS PROGRAMAS DE PREGRADO OFRECIDOS
POR LA UNIVERSIDAD NACIONAL ABIERTA Y A DISTANCIA UNAD**

**FEDERICO CRUZ LADINO
PAULA ANDREA SUÁREZ GAVIRIA**

**UNIVERSIDAD NACIONAL ABIERTA Y A DISTANCIA – UNAD
ESCUELA DE CIENCIAS BÁSICAS, TECNOLOGÍA E INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
MEDELLIN
2020**

**DISEÑO DE EXPERIMENTOS APLICADO AL ANÁLISIS DE LA
DINÁMICA DEL FRAUDE ACADÉMICO EN LOS PROGRAMAS DE
PREGRADO OFRECIDOS POR LA UNIVERSIDAD NACIONAL ABIERTA
Y A DISTANCIA UNAD**

**FEDERICO CRUZ LADINO
PAULA ANDREA SUÁREZ GAVIRIA**

**DIRECTOR:
GABRIEL JAIME RIVERA LEÓN
M. SC. EN INGENIERÍA ADMINISTRATIVA**

**UNIVERSIDAD NACIONAL ABIERTA Y A DISTANCIA – UNAD
ESCUELA DE CIENCIAS BÁSICAS, TECNOLOGÍA E INGENIERÍA
PROGRAMA DE INGENIERÍA INDUSTRIAL
MEDELLIN
2020**

Tabla de Contenido

Resumen	5
Introducción.....	6
1. Planteamiento del problema	7
2. Justificación	8
3. Objetivos.....	9
3.1 Objetivo general	9
3.2 Objetivos específicos	9
4. Marco Teórico.....	9
4.5. Diseño de experimentos	177
4.6. Diseño de experimentos como herramienta para validar modelos de simulación	22
4. Metodología.....	233
5. Modelo de simulación basada en agentes.....	24
6. Modelo de diseño de experimentos	31
7. Resultados.....	34
8. Conclusiones	49
9. Recomendaciones	51
10. Referencias	52

Tabla de ilustraciones

Ilustración 1. Proceso de modelado con abstracción intermedia.....	12
Ilustración 2.En los modelos formales construidos mediante simulación basada en agentes, los componentes básicos del sistema real y las interacciones entre ellos están explícita e individualmente representados en el modelo.	15
Ilustración 3.Relaciones del mundo real y la simulación mundial con la validación.	16
Ilustración 4. Artículos publicados a lo largo de los años.	18
Ilustración 5.Familia de diseños factoriales 2^k ($k \leq 5$).	20
Ilustración 6. Diseños factoriales 2^k y sus efectos de interés, $k \leq 5$	22
Ilustración 7.Modelo de simulación basada en agentes	24
Ilustración 8. Funcionamiento del modelo de simulación basada en agentes.	25

Tabla de análisis

Tabla 1. Factores y niveles del diseño experimental propuesto	27
Tabla 2.Matriz de diseño en orden aleatorio (unidades codificadas).....	29
Tabla 3. ANOVA para el modelo de fraude académico.....	32
Tabla 4. ANOVA para el modelo ajustado.....	41

Tabla de gráficas

Gráfica 1.Diagrama de Pareto de efectos estandarizados	29
Gráfica 2. Gráfica de efectos normales estandarizados.....	30
Gráfica 3. Efectos principales para Y: factor A.....	33
Gráfica 4. Efectos principales para Y: factor B.....	33
Gráfica 5.Efectos principales para Y: factor C.....	34
Gráfica 6.Efectos principales para Y: factor D.....	35
Gráfica 7.Efectos principales para Y: factor E.....	36
Gráfica 8.Efecto interacción para Y: interacción AE.....	36
Gráfica 9.Gráfico de probabilidad normal de los residuos estandarizados.....	38
Gráfica 10.Gráfico de los residuos vs predichos	38
Gráfica 11.Gráfico de los residuos estandarizados vs orden.	39
Gráfica 12.Diagrama de Pareto de efectos estandarizados modelo ajustado	40
Gráfica 13.Gráfica de efectos normales estandarizados del modelo ajustado	40
Gráfica 14. Gráfico de probabilidad normal de los residuos estandarizados del modelo ajustado.	42
Gráfica 15. Gráfico de los residuos vs predichos del modelo ajustado.	42
Gráfica 16.Gráfico de los residuos estandarizados vs orden del modelo ajustado	43

Resumen

En este estudio, se propone aplicar la metodología de diseño de experimentos para validar el modelo de simulación basada en agentes construido en la fase anterior del proyecto en el cual se enmarca esta propuesta de investigación. Adicionalmente, se propone usar el diseño de experimentos para evaluar el efecto del cambio en los niveles de los factores seleccionados en el diseño experimental, sobre el número de casos de fraude detectados en los programas de pregrado de la UNAD. Esta investigación, surge como una forma de mostrar que el modelo de simulación representa la dinámica real del fenómeno del fraude en los programas de pregrado de la UNAD, y como una forma de evaluar el efecto de diferentes políticas sobre el número de casos de fraude detectado. A partir de los resultados obtenidos por el modelo de diseño de experimentos, tenemos que, los factores más importantes a la hora de determinar el número de estudiantes que cometen fraude son, en primer lugar, la percepción que tenga el estudiante sobre qué tan laxo o que tan estricto frente a las conductas tipificadas como fraude son los docentes; también tenemos que, el número de actividades que se le asignen durante un período académico a los estudiantes en un curso, y la cantidad de estudiantes con los que un estudiante tenga contacto en ese período académico, va a determinar si el estudiante comete fraude en sus actividades académicas o no.

Introducción

Este trabajo está motivado en el diseño de experimentos aplicado al análisis de la dinámica del fraude académico en los programas de pregrado ofrecidos por la Universidad Nacional Abierta y a Distancia UNAD, al encontrar las diferentes dificultades que se evidenciaron en el ámbito educacional, en el momento en que los estudiantes presenten sus trabajos escritos. Para ello se recurre a la técnica de diseño de experimentos 2k como herramienta para validar un modelo de simulación, cuyo proceso al ser implementado, permite analizar y sacar conclusiones objetivas de lo que en este caso se quiere encontrar.

Las características principales de este tipo de investigación es abordar la problemática desde un diseño experimental del tipo Diseño Factorial 25, el cual, se apoya en un diseño factorial en el que se tienen una variable respuesta, cinco factores y dos niveles para cada uno de estos factores.

Para analizar la problemática de la dinámica del fraude es necesario mencionar las cinco causas encontradas para la realización de ésta investigación, para ello a continuación se nombran en el orden respectivo: número de actividades por asignatura, percepción de fraude del docente, percepción del estudiante frente a la dificultad de la actividad, percepción del estudiante frente a severidad de sanciones, tamaño de la red del estudiante (número de compañeros con los que tiene contacto el estudiante).

Por otra parte éste trabajo tiene como interés académico aplicar la técnica de simulación para explicar la dinámica del fraude con el fin de encontrar resultados concretos que permitan que las universidades apliquen sus políticas institucionales durante el proceso pedagógico de los estudiantes.

Para dar inicio a la investigación de campo se partió desde la realización de una encuesta realizada a los estudiantes de la Escuela de Ciencias Básicas, Tecnología e Ingeniería (ECBTI) de la Universidad Nacional Abierta y a Distancia (UNAD), cuyas preguntas y respuestas sirvieron como punto clave para erigir el modelo de simulación basada en agentes en donde se utilizó el software NetLogo, el cual permite identificar si el estudiante incurre o no en el fraude. Con los hechos realizados anteriormente se da continuidad a aplicar el diseño de experimento para poder validar los resultados obtenidos por el modelo de simulación.

La finalidad de ésta investigación en la que se quiere analizar la dinámica del fraude académico, es validar el modelo de simulación basado en agentes por medio del diseño de experimentos que se propuso, para con el resultado evaluar el impacto que conlleva incurrir en esta falta e identificar algunas de las variables más importantes que afectan el número de casos de plagio que se puedan presentar, y con esto, plantear estrategias que permitan mitigar o acabar con el número de casos de fraude que se presentan en los programas ofrecidos por la universidad.

1. Planteamiento del problema

A nivel de educación Universitaria, se han realizado estudios sobre el fraude académico, encontrándose que un alto porcentaje de los estudiantes admiten que en algún momento han realizado fraude durante la presentación de exámenes y/o trabajos. No obstante, y aunque las universidades tienen normas que penalizan este tipo de conductas, sin llegar a ser una fuerza persuasiva para la prevención de este problema (Diaz, Gonzalez, & Carmona, 2010). Debido a esto, es necesario analizar la dinámica del fraude académico en las instituciones universitarias como una forma de identificar las variables que influyen en éste problema y así plantear estrategias que permitan persuadir a los estudiantes de los efectos desfavorables de estas actividades, tanto para él como para el profesional que quiere lograr ser??.

Aunque el estudio de la dinámica del fraude académico en instituciones de educación superior, no es un tema nuevo, ya que se han reportado resultados de investigaciones realizadas tanto en Colombia como en otros países, desde la década de 1950, la mayoría de estos estudios sólo abordan el problema desde su caracterización. Para abordar este tipo de problemas, se aplican técnicas de muestreo para analizar el fenómeno y su dinámica, como una forma de explicar el fenómeno, sus causas y consecuencias. En ninguno de los estudios encontrados se trata de abordar el problema desde la simulación, cómo una forma de construir una representación del sistema que sirva como una herramienta para evaluar políticas y su efectividad. La presente investigación busca además de aplicar técnicas de muestreo para analizar el fenómeno y conocer su dinámica, construir un modelo de simulación que sirva como herramienta para evaluar la efectividad de las políticas que en materia de fraude implementen las instituciones objeto de estudio.

Con este modelo, se busca contribuir al mejoramiento de la calidad de formación pedagógica institucional de las instituciones participantes en el estudio. Lo cual permitirá garantizar la transparencia de los procesos pedagógicos planteados.

A partir de los anterior, nos planteamos la siguiente pregunta de investigación:

¿Podrá las técnicas de simulación y diseño de experimentos, explicar el fraude académico relacionado con los programas universitarios virtuales en relación a la normativa universitaria vigente?

2. Justificación

Aunque en la literatura se encuentra diversas definiciones sobre el “fraude” así mismo como estudios sobre el tema, abordado desde diferentes disciplinas. Para el presente trabajo de investigación, se parte de la definición del “fraude” como acto de mala fé, con el que se busca engañar a alguien para obtener una ventaja o un lucro indebido (Hilber, 1987). Mientras que el “fraude académico” se define como el evento que se presenta, cuando una persona se apropia de forma indebida del conocimiento, ideas o descubrimientos realizados por alguien más, con el objetivo de obtener un beneficio personal (Muhney, y otros, 2008).

El fraude académico es una práctica que, aunque es muy frecuente en los estudiantes universitarios tanto presencial como virtual, no deja de ser una práctica deshonesta que puede desencadenar malas prácticas en los profesionales. De ahí la importancia de generar estrategias para mitigar su presencia en las instituciones de educación superior del país. Aunque se han realizado estudios sobre el fraude académico desde la década de 1950, estos estudios se han centrado sólo en la descripción del fenómeno, usando técnicas de muestreo. En ninguno de estos estudios, se ha abordado el problema desde la construcción de un modelo de simulación que sirva de apoyo a la toma de decisiones. Es aquí en donde nuestra investigación, puede generar un gran aporte a la búsqueda de soluciones que permitan disminuir la ocurrencia de fraude en las instituciones de educación superior.

Para abordar el problema, se tendrá como referencia a los programas académicos de la Universidad Nacional Abierta y Distancia, debido a que todos sus programas son virtuales y a los programas virtuales de la Facultad de Ingeniería de la Universidad de Antioquia. Para lo cual, se aplicará una encuesta a los estudiantes, en donde estos de forma anónima respondan un cuestionario, que nos permita conocer el estado actual del problema en dichos programas. A partir de esta encuesta, se construirá un modelo de simulación, que represente la dinámica del fraude académico en los programas virtuales de las instituciones involucradas. Con este modelo de simulación se evaluarán políticas con el objetivo de disminuir la incidencia del fraude académico en los programas evaluados. Adicionalmente, se plantearán y aplicarán diseños experimentales al modelo de simulación, como una forma de validar los resultados obtenidos por este.

Los resultados de este estudio, les permitirán a las universidades participantes, analizar la efectividad de las políticas y la normatividad que aplican en la actualidad para el control del fraude académico, en los programas objeto de estudio. Adicionalmente, pueden utilizarse para evaluar nuevas políticas que les permitan disminuir la ocurrencia del fraude en dichos programas. Estos resultados, también pueden ser aprovechados por todas aquellas universidades del país que ofrecen programas académicos virtuales.

3. Objetivos

3.1 Objetivo general

Analizar la dinámica del fraude académico en los programas universitarios virtuales, ofrecidos por la Universidad Nacional Abierta y Distancia usando técnicas de Diseño de Experimentos.

3.2 Objetivos específicos

- Proponer un diseño experimental aplicable al modelo de simulación basada en agentes desarrollado, que permita analizar la dinámica del fraude académico en los programas de pregrado ofrecidos por la UNAD.
- Validar el modelo de simulación basada en agentes por medio del diseño experimental propuesto.
- Evaluar, usando el diseño de experimentos, el impacto de diferentes políticas sobre el número de casos de fraude detectados por período académico.

4. Marco Teórico

4.1. Definición de Plagio

Para analizar en detalle el término y las diferentes situaciones en las que se reincide en el plagio, se enuncia a continuación los puntos de vista de varios investigadores:

(Aguilar Passano, 2019) Puntualiza un poco de historia con el origen latino del sustantivo plagio (plagium) se remonta al siglo III aC y su acepción era “usar al esclavo ajeno como si fuera propio”. Plagium, a su vez, proviene del griego (plagios) y se utilizaba para significar un “secuestro”.

En un informe (Soto Rodríguez , 2012) considera que:

- El plagio ocurre cuando se toman ideas o palabras escritas por otros sin reconocer de forma directa el haberlo hecho. (Girón, 2008, p. 6)
- Se produce también al presentar como propio un trabajo de forma parcial o total sin ser el autor o autora de dicho trabajo. (Girón, 2008, p. 6)
- Al actuar de mala fe deliberadamente al copiar la propiedad intelectual de otros para producir un daño a los autores originales. (Girón, 2008, p. 6)
- Se considera que se comete plagio al copiar cualquier objeto de fondo o de forma, ya sea una situación, un desarrollo o incluso una simple frase. (Girón, 2008, p. 6)
- Inclusive se comete plagio al copiar lo dicho por otro en un discurso o dictado sin hacer referencia a la persona que lo dijo. (Girón, 2008, p. 7)

- Al imitar un modelo y reproducirlo de forma idéntica de nuevo se incurre en este delito. (Girón, 2008, p. 7)

Con lo anterior dicho por (Soto Rodríguez , 2012) se percibe el problema que abarca la gran deshonestidad que se da al presentar trabajos de investigación en donde se copia y pega las publicaciones científica, atentando contra los derechos del autor.

En el ámbito jurídico, plagiar conlleva, según Girón Castro en un informe de la Universidad Sergio Arboleda, dos clases de delitos: en primer lugar, “usar ideas, información o expresiones de otra persona sin darle el debido reconocimiento”, constituye robo de propiedad intelectual, y, en segundo lugar, “hacer pasar las ideas, información o expresiones de otra persona como si fueran propias para obtener buenas calificaciones u otras ventajas”, constituye fraude (Aguilar Passano, 2019).

En Colombia de acuerdo con las investigaciones y su sistema de creencias, la Constitución Política de 1991 en diferentes artículos y específicamente en el artículo 61, dice: “el Estado protegerá la propiedad intelectual por el tiempo y mediante las formalidades que establezca la ley”. (Sanabria, 2014)

Consecuente, el artículo 61 con el artículo 58 advierte: “se garantiza la propiedad privada y los demás derechos adquiridos con arreglo a las leyes civiles”. “La propiedad es una función social que implica obligaciones”. (Sanabria, 2014)

El Código Penal en el artículo 270 se refiere a la violación de los derechos morales del autor con penas aumentadas por el artículo 14 de la Ley 890 de 2004 que implica prisión de 32 a 90 meses y multa de 26,66 a 300 salarios mínimos legales mensuales vigentes. (Sanabria, 2014).

Esto, a su vez significa lo que se espera por parte del adoctrinamiento del plagio; brindando protección y manteniendo la originalidad de los escritos, además del fin al que se pretende llegar con cuidar lo que se conoce, respetando la propiedad intelectual y obligar a los investigadores a analizar y crear sin incurrir en la ilegalidad.

(Rojas Chavarro, 2010) Afirma, que en el medio académico, es común este tipo de infracción, en la elaboración de trabajos, proyectos o tesis de grado. Ésta, encuentra un escenario favorable debido a la facilidad que presenta el entorno digital para copiar y pegar un texto, por la disponibilidad de información y facilidad de búsqueda en Internet y por el hecho de que los profesores orientan los trabajos de sus estudiantes hacia la recolección y presentación de información, más que hacia la solución de problemas o el desarrollo de su capacidad creadora.

En conclusión con lo anterior, el plagio, en la actualidad y con la tecnología digital que se cuenta, en el sistema académico implica muchos problemas; pues en la búsqueda de la solución se invita a los docentes y estudiantes que se percaten de la ética que se debe tener al entregar trabajos, pues es en estos ambientes donde más se evidencian este tipo de comportamientos, ya que por parte del alumnado

solo se busca la vía fácil; por esto las instituciones se ven cada vez más comprometidas en legitimar culturas de comportamientos conscientes y capacidad creadora, en donde se acabe con ésta comodidad.

La honestidad debe ser la premisa que guíe las actuaciones del individuo y lograr que esto se dé, es una tarea que va desde que nacemos hasta la tumba (García Márquez, 1994).

4.2. Simulación de sistemas

Para empezar a hablar de la Simulación de Sistemas, es importante previamente conocer estas palabras por separado. Por consiguiente, la simulación Tomas H. Naylor la define así:

Simulación es una técnica numérica para conducir experimentos en una computadora digital. Estos experimentos comprenden ciertos tipos de relaciones matemáticas y lógicas, las cuales son necesarias para describir el comportamiento y la estructura de sistemas complejos del mundo real a través de largos períodos de tiempo. (Bu, 2003)

Dentro de las ventajas más comunes que ofrece la simulación (Eduardo García Dunna, Heriberto García Reyes, Leopoldo Eduardo Cárdenas Barrón, 2006) mencionan las siguientes:

- a. Es muy buena herramienta para conocer el impacto de los cambios en los procesos sin necesidad de llevarlos a cabo en la realidad.
- b. Mejora el conocimiento del proceso actual al permitir que el analista vea cómo se comporta el modelo generado bajo diferentes escenarios.
- c. Puede utilizarse como medio de capacitación para la toma de decisiones.
- d. Es más económico realizar un estudio de simulación que hacer muchos cambios en los procesos reales.
- e. En problemas de gran complejidad, la simulación permite generar una buena solución.
- f. En la actualidad los paquetes de software para simulación tienden a ser más sencillos, lo que facilita su aplicación.
- g. Gracias a las herramientas de animación que forman parte de muchos de esos paquetes es posible ver cómo se comportará un proceso una vez que sea mejorado.

De manera paralela a la simulación, es importante también profundizar en el concepto de sistemas, que según Bertalanffy, 1968, considera básicamente se puede definirse como un «conjunto de elementos en interacción» (IZQUIERDO, GALÁN, SANTOS, & DEL OLMO, 2008)

Para llevar a cabo un modelo, es muy importante los términos nombrados anteriormente, pues con la unión de ellos se hace posible el desarrollo y perfeccionamiento de un problema de manera más técnica. En una investigación para la revista IMPERIA, en donde se refieren a la simulación de sistemas, haciendo una pequeña comparación de como anteriormente se solucionaban modelos matemáticos y como llegan los ordenadores a realizar la misma tarea de simulación de sistemas complejos pero de una manera más asertiva y rápida. De éste modo, él resume, “un modelo computacional es un modelo formal (que por lo tanto puede expresarse en lenguaje matemático como un conjunto de ecuaciones), y la simulación computacional es una herramienta que nos permite estudiarlo más allá de los límites actuales de las matemáticas”. (IZQUIERDO, GALÁN, SANTOS, & DEL OLMO, 2008)

(IZQUIERDO, GALÁN, SANTOS, & DEL OLMO, 2008) Nos muestra en la siguiente figura 1, de cuando tratamos de modelizar sistemas complejos, raramente construimos un modelo formal directamente sobre el sistema real. Normalmente la mayoría de los científicos abstraen, consciente o inconscientemente, un modelo intermedio de cómo se cree que el sistema real funciona, y es esta abstracción —a menudo incompleta y no formal— la que modelizamos formalmente.

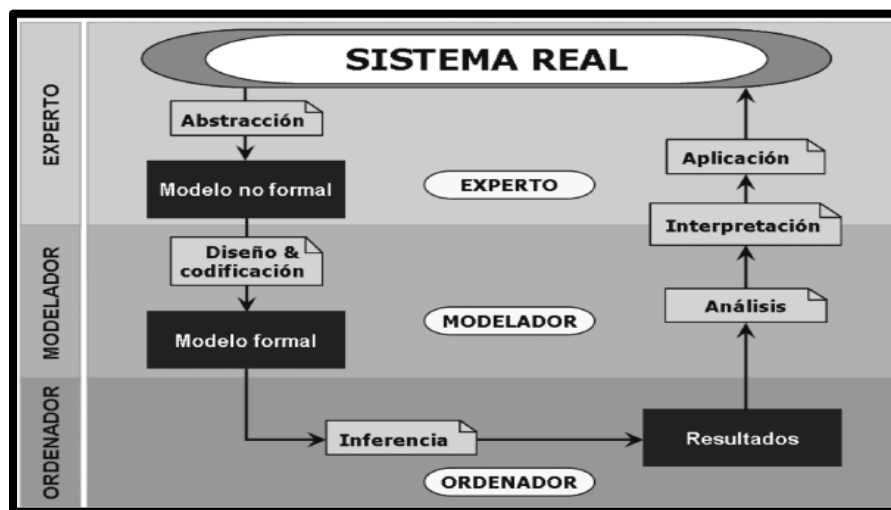


Ilustración 1. Proceso de modelado con abstracción intermedia

Fuente: (IZQUIERDO, GALÁN, SANTOS, & DEL OLMO, 2008)

Para explicar cada una de las etapas del proceso de modelado de sistemas complejos encontramos útil distinguir tres roles (Drogoul et al., 2003): *experto*, *modelador* y *ordenador*.

El *experto* es una persona con un gran conocimiento del sistema real que se pretende modelizar y de su funcionamiento. Sin embargo, el experto no está necesariamente familiarizado con la creación y el uso de modelos formales. El *modelador*, por el contrario, es un profesional cuya labor principal consiste en

diseñar, implementar y analizar modelos formales. El modelador tiene, por tanto, conocimientos avanzados de algún formalismo que permita crear y estudiar modelos formales (p. ej. cálculo simbólico, lógica proposicional, o un lenguaje de programación). Finalmente, el *ordenador* es el encargado de ejecutar o resolver el modelo formal, i.e. el ordenador deduce las implicaciones lógicas que se derivan de las premisas que definen el modelo y de las condiciones iniciales.

(Miquel Àngel Piera, 2006) Considera también tener en cuenta las siguientes etapas del proceso, cuyos conceptos fueron asociados a él presente proyecto de investigación:

Recogida de datos: En esta fase se realiza la investigación de la cantidad de estudiantes que incurren en alguna conducta considerada como fraude, es aquí donde el experto sale al campo de estudio y hace su labor de búsqueda de información, obteniendo datos del número de actividades por asignatura, percepción de fraude del docente, percepción del estudiante frente a la dificultad de la actividad, percepción del estudiante frente a severidad de sanciones, tamaño de la red del estudiante (estudiantes que entre si se comunican).

Construcción del modelo: De acuerdo con la información obtenida por el experto en la etapa anterior, se procede al diseño del experimento, en donde éste plantea un hipótesis de interés, en este caso es la cantidad de estudiantes que incurren en alguna conducta considerada como fraude, aquí se quiere investigar la influencia de 5 factores que son, número de actividades por asignatura, percepción de fraude del docente, percepción del estudiante frente a la dificultad de la actividad, percepción del estudiante frente a severidad de sanciones, tamaño de la red del estudiante (estudiantes que entre si se comunican), estas poseen 2 variables de respuestas, todo esto y teniendo la comprensión del problema se procede a convertirlo en un Diseño Factorial 2^5 .

Verificación y Validación: Se comprueba que el modelo Diseño Factorial 2^5 se comporte como es de esperar al ejecutar la simulación, y a partir de esto obtener los resultados; en donde debe existir una correspondencia entre el sistema real y el modelo al hacer la interpretación de éstos.

Implementación: Con los resultados arrojados por la simulación, lo que se pretende es aplicarlos en el sistema real para llegar a la mejor decisión en la solución del problema planteado en la investigación.

4.3. Simulación basada en agentes

(T. Echaveguren, 2016) Define el concepto de la simulación basada en agentes como una herramienta que se utiliza preferentemente para modelar sistemas adaptativos, emergentes, cerrados o abiertos en donde existe una interacción entre las decisiones técnicas y los agentes que toman decisiones en respuesta a esas decisiones técnicas. De acuerdo a (Zhang y Li, 2010), un modelo basado en agentes

requiere de un ambiente en el cual se desenvuelven los agentes, de suerte que la proyección de las decisiones e interacciones entre agentes se proyecta en dicho ambiente.

Para (Millán Arroyo Menéndez & Samer Hassan Collado, 2007) la Simulación Basada en Agentes, a pesar de su reciente desarrollo, está teniendo un rápido crecimiento. Esto es debido principalmente al potencial que ofrece en particular en los dominios caracterizados por un alto grado de localización (como en las redes sociales, donde los nodos y enlaces son esenciales) y distribución (no centralizado y dirigido por una jerarquía, sino distribuido horizontalmente, aunque es posible introducir capas de decisión), frente al modelado basado en ecuaciones, más apropiado para sistemas centralizados dominados más por leyes físicas que por procesos de información.

De acuerdo con lo anterior (José Ignacio García & Valdecasas Medina, 2011) ayuda a entender la simulación basada en agentes aclarando que son experimentos virtuales, que son los agentes y cuál es su entorno.

Un experimento consiste en aplicar algún tratamiento a una parte aislada de la realidad y observar qué ocurre; posteriormente, el objeto de análisis que ha sido tratado es comparado con otro objeto equivalente que no ha recibido ningún tratamiento (llamado «control»). La gran ventaja de la experimentación es que permite asegurar que el tratamiento aplicado al objeto de estudio es de hecho la causa de los efectos observados, ya que solo el tratamiento es lo que difiere entre el objeto de análisis y el control (Gilbert, 2008: 3).

Los agentes virtuales, son en realidad módulos o partes de los programas, representan a los actores reales. Existe, pues, una equivalencia entre los agentes del mundo virtual y los actores del mundo real, hecho que facilita tanto el diseño como la interpretación de los resultados de los experimentos virtuales frente a otros tipos de simulaciones (Gilbert, 2008: 14).

El entorno virtual, que representa al medio real en el que los actores reales operan, es el hábitat donde los agentes virtuales interactúan entre sí. Dicho entorno puede representar espacios geográficos con barreras físicas o fuentes de energía, por ejemplo, donde los agentes tienen coordenadas que indican su localización, aunque en otros modelos puede también representar otros tipos de espacios, como por ejemplo espacios de conocimiento (Gilbert et al., 2001) o de relaciones (redes).

Según (IZQUIERDO, GALÁN, SANTOS, & DEL OLMO, 2008) lo que distingue a la simulación basada en agentes de otras técnicas de modelado es la forma en que se construye la primera abstracción del sistema real y, consecuentemente, el modelo formal (ver fig. 1). En los modelos formales construidos mediante simulación basada en agentes, los componentes básicos del sistema real están explícita e individualmente representados en el modelo (Edmonds et al., 2001). De esta forma, como se muestra en la figura 2, las fronteras que definen a los componentes básicos

del sistema real se corresponden con las fronteras que definen a los agentes del modelo, y las interacciones que tienen lugar entre los componentes básicos del sistema real se corresponden con las interacciones que tienen lugar entre los agentes del modelo (Edmonds et al., 2001; Galán et al., 2008). Esta correspondencia directa contrasta con el tradicional uso de 'agentes representativos' y es capaz de aumentar el realismo y el rigor científico de los modelos formales así construidos.

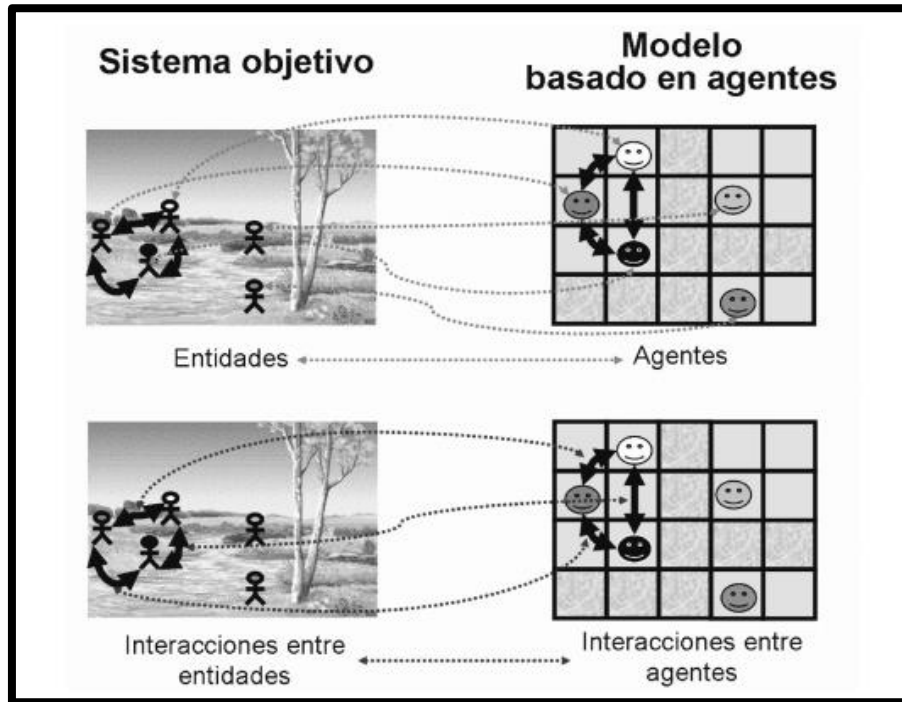


Ilustración 2. En los modelos formales construidos mediante simulación basada en agentes, los componentes básicos del sistema real y las interacciones entre ellos están explícita e individualmente representados en el modelo.

Fuente: (IZQUIERDO, GALÁN, SANTOS, & DEL OLMO, 2008)

4.4. Validación de un modelo de simulación

Los modelos de simulación son utilizados cada vez más para resolver problemas y ayudar en la toma de decisiones. Los desarrolladores y usuarios de estos modelos, los responsables de la toma de decisiones que utilizan la información obtenida de los resultados de estos modelos y las personas afectadas por las decisiones basadas en dichos modelos se preocupan por si un modelo y sus resultados son correctos. Esta preocupación se aborda a través de la validación del modelo. La validación del modelo generalmente se define como la "justificación de que un modelo computarizado dentro de su dominio de aplicabilidad posee un rango satisfactorio de precisión consistente con la aplicación prevista del modelo" (Schlesinger, 1979) y es la definición en la cual nos fundamentamos.

Por lo general, se requieren numerosos conjuntos de condiciones experimentales para definir el dominio de la aplicabilidad prevista de un modelo. Un modelo puede ser válido para un conjunto de condiciones experimentales y no válido en otro. Un modelo se considera válido para un conjunto de condiciones experimentales si la precisión del modelo está dentro de su rango aceptable, que es la cantidad de precisión requerida para el propósito previsto del modelo. Esto generalmente requiere identificar las variables de salida de interés del modelo y especificar el rango de precisión aceptable requerido para cada variable (Sargent, 2011).

Según (José Ignacio García & Valdecasas Medina, 2011) un modelo es válido si al ingresar toda la información al simulador, éste arroja los resultados con éxito. Ésta validación se compara con el problema real en donde no se espera que se encuentren resultados exactos ya que el modelo y la realidad son procesos con datos estocásticos.

En la etapa de validación (Manish Kumar , Andrew S. Whittaker, 2018) dice que se debe cumplir con los siguientes tres condiciones:

1. El modelo está claramente definido, incluida la realidad de interés, su uso revisado y las cantidades de respuesta de interés.
2. Se utiliza un modelo computacional verificado.
3. Las incertidumbres en los resultados experimentales se cuantifican.

A continuación (Sargent, 2010) muestra el proceso de validación de un modelo de simulación y su relación con el sistema real:

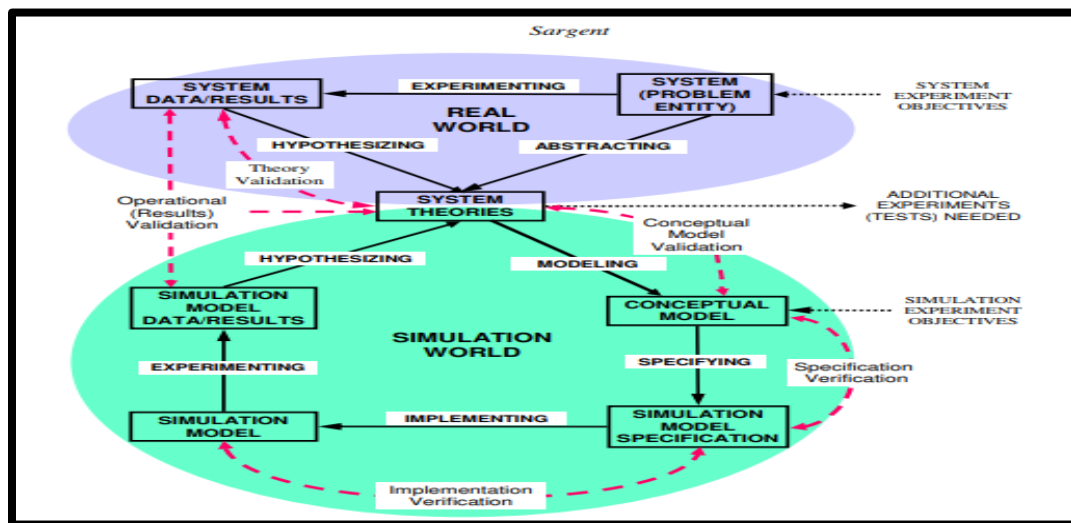


Ilustración 3. Relaciones del mundo real y la simulación mundial con la validación.

Fuente: (Sargent, 2010)

En la figura 3, la primera validación conceptual del modelo, se aplica en el mundo real, en donde, eligiendo un campo de investigación, se experimenta o se busca una dificultad; a partir de ésta, se formula una hipótesis y se adquieren datos y resultados teóricos, ésta información debe ser validada y para ello debe ser comparada con las causas del problema, que son los datos y resultados que muestra la teoría del sistema.

En un segundo plano vemos el mundo de la simulación en donde encontramos, el modelo de simulación, el cual es diseñado por el experto a partir del modelo conceptual, éste debe ser adaptado a las especificaciones del software. La Validación del modelo computarizado se da al obtener los resultados y compararlos con los del sistema real, aquí debe existir una similitud la cual garantiza la confiabilidad de dicho proceso.

Una vez se tenga éxito en los datos arrojados por el software, se procede a la validación operacional, en donde se hace un estudio más riguroso de la salida de estos resultados y se determina que en realidad son exactos al propósito de estudio.

Por último y en el mismo orden de importancia se encuentra la validación de la implementación la cual asegurar que el modelo de simulación se haya implementado de acuerdo con la especificación del modelo de simulación. (Sargent, 2010)

4.5. Diseño de experimentos

El diseño de experimentos se refiere al proceso de planear, diseñar y analizar el experimento para que se puedan sacar conclusiones válidas y objetivas de manera efectiva y eficiente. Para sacar conclusiones estadísticamente sólidas del experimento es necesario integrar simples y fuertes métodos estadísticos dentro de la metodología de diseño de experimentos. El éxito de cualquier experimento diseñado industrialmente depende de una buena planificación, elección adecuada del diseño, análisis estadístico de datos y habilidades de trabajo en equipo.

El diseño de experimentos es la aplicación del método científico para forjar conocimiento acerca de un sistema o proceso, por medio de pruebas proyectadas adecuadamente. Esta metodología se ha ido fortaleciendo como un conjunto de técnicas estadísticas y de ingeniería, que permiten comprender mejor contextos complejos de relación causa-efecto (Gutierrez & de la Vara, 2008).

Para que un estudio experimental sea exitoso es necesario realizar, por etapas, diferentes actividades.

Planeación: Se deben hacer investigaciones preliminares que conduzcan a entender y delimitar el estudio, elegir las variables que mejor reflejen al objeto de

estudio, incluir toda la información disponible, determinar las repeticiones que se harán para cada tratamiento, organizar la forma operativa y ejecutar lo planeado.

Análisis: Se debe recurrir a métodos estadísticos inferenciales para ver si las diferencias o efectos muestrales son lo suficientemente grandes para que garanticen diferencias poblacionales.

Conclusión: Se decide qué medidas implementar para generalizar el resultado del estudio y para garantizar que las mejoras se mantengan.

Los desarrollos revolucionarios en los campos asociados a la computación y la tecnología han motivado tanto al personal académico como al personal industrial a optar por experimentos informáticos. Esto inspiró a varios investigadores a trabajar en el diseño de experimentos informáticos y sus aplicaciones, lo cual es evidente en la Figura 3, la cual muestra el número de artículos publicados en el campo del diseño de experimentos a lo largo de los años. Aunque la investigación más temprana en el campo del diseño moderno de experimentos se remonta a fines de la década de 1940, atrajo una atención significativa a fines de la década de 1970. Por lo tanto, la Figura 4 abarca desde 1970 hasta 2016, a partir de entonces, el campo está en auge, lo que es evidente por la fuerte tendencia de crecimiento que se evidencia (Garud, Karimi, & Kraft, 2017).

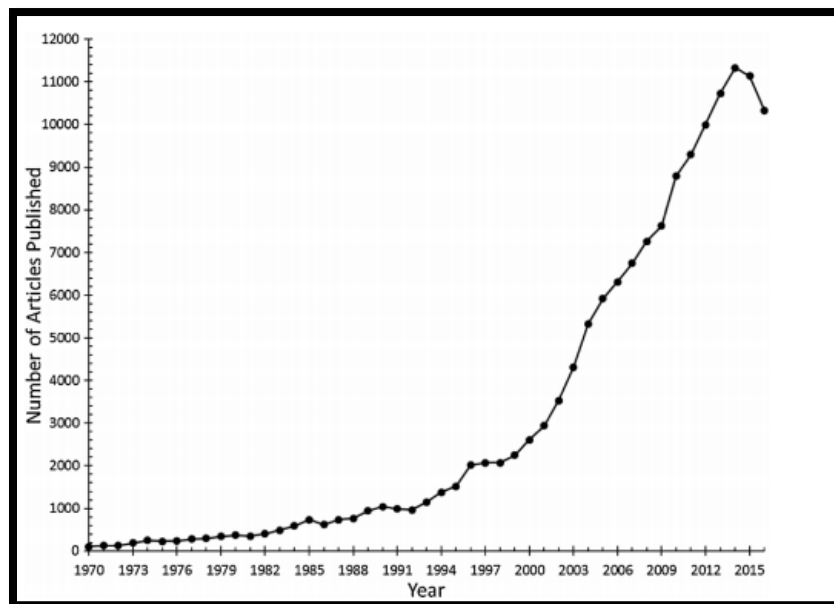


Ilustración 4. Artículos publicados a lo largo de los años.

Fuente: tomado de Garud, Karimi, & Kraft, (2017).

DISEÑOS FACTORIALES

La función principal de un diseño factorial es estudiar la consecuencia de varios factores sobre una o varias alternativas, cuando se tiene la misma disposición sobre todos los factores. Los factores pueden ser de tipo cualitativo o de tipo cuantitativo. Para analizar la manera en que interviene cada factor sobre la variable de respuesta es conveniente designar al menos dos niveles de prueba para cada uno de ellos. Con el diseño factorial completo, se corren aleatoriamente todas las probables combinaciones que se puedan formar con los niveles de los factores a intervenir (Gutierrez & de la Vara, 2008).

EXPERIMENTOS DE UN FACTOR

Muchos análisis de un factor se hacen con base en el diseño completamente al azar (DCA), que es el más simple de todos los diseños que se utilizan para este tipo de casos, dado que sólo consideran dos fuentes de variabilidad: los tratamientos y el error aleatorio. Este diseño debe su nombre a la cantidad de las corridas experimentales que se realizan en orden aleatorio completo, de esta manera, si durante el estudio se hacen en total N pruebas, éstas se corren aleatoriamente, de manera que los posibles efectos ambientales y temporales se vayan repartiendo equitativamente entre los tratamientos (Gutierrez & de la Vara, 2008).

El análisis de varianza o ANOVA es la técnica central en el análisis de datos experimentales cuya idea fundamental es separar la variación total en las partes con las que contribuye cada fuente de variación en el experimento. En el caso del diseño completamente al azar se separan la variabilidad debida a los tratamientos y la debida al error. Cuando la primera predomina notoriamente sobre la segunda, es cuando se concluye que los tratamientos tienen efecto o en otras palabras las medias son diferentes. Cuando los tratamientos no dominan contribuyen igual o menos que el error, por lo que se concluye que las medias son iguales (Gutierrez & de la Vara, 2008).

DISEÑO DE EXPERIMENTOS 2^k

El conjunto de diseños factoriales completos 2^k es uno de los de mayor impacto en la industria y en la investigación, debido a su eficacia y versatilidad, estos son útiles principalmente cuando el número de factores a estudiar está entre dos y cinco, rango en el cual su tamaño se encuentra entre cuatro y 32 tratamientos; lo cual es una cantidad manejable en muchas situaciones experimentales. Si el número de factores es mayor que cinco se recomienda utilizar un diseño factorial fraccionado 2^{k-p} . En general, los factoriales en dos niveles, sean completos o fraccionados, constituyen el conjunto de diseños de mayor impacto en las aplicaciones (Gutierrez & de la Vara, 2008).

DISEÑO FACTORIAL GENERAL 2^k

En el caso general del diseño 2^k se consideran k factores con dos niveles cada uno, y tiene $2k$ tratamientos o puntos de diseño. Las k columnas y $2k$ renglones que conforman la matriz para este diseño, considerando una réplica, se construyen así:

- En la primera columna, que corresponde a los niveles del factor A, se alternan signos + y -, empezando con - hasta llegar a los $2k$ renglones;
- En la segunda columna se alternan dos signos menos con dos signos más;
- En la tercera, se alternan cuatro signos menos y cuatro signos más, y así sucesivamente hasta la k -ésima columna compuesta por $2k - 1$ signos -, seguidos de $k - 1$ signos +.

Tratamiento	Notación de Yates	A	B	C	D	E	Tratamiento	Notación de Yates	A	B	C	D	E
1	(1)	-	-	-	-	-	17	e	-	-	-	-	+
2	a	+	-	-	-	-	18	ae	+	-	-	-	+
3	b	-	+	-	-	-	19	be	-	+	-	-	+
4	ab	+	+	-	-	-	20	abe	+	+	-	-	+
5	c	-	-	+	-	-	21	ce	-	-	+	-	+
6	ac	+	-	+	-	-	22	ace	+	-	+	-	+
7	bc	-	+	+	-	-	23	bce	-	+	+	-	+
8	abc	+	+	+	-	-	24	abce	+	+	+	-	+
9	d	-	-	-	+	-	25	de	-	-	-	+	+
10	ad	+	-	-	+	-	26	ade	+	-	-	+	+
11	bd	-	+	-	+	-	27	bde	-	+	-	+	+
12	abd	+	+	-	+	-	28	abde	+	+	-	+	+
13	cd	-	-	+	+	-	29	cde	-	-	+	+	+
14	acd	+	-	+	+	-	30	acde	+	-	+	+	+
15	bcd	-	+	+	+	-	31	bcde	-	+	+	+	+
16	abcd	+	+	+	+	-	32	abcde	+	+	+	+	+

Ilustración 5. Familia de diseños factoriales 2^k ($k \leq 5$).

Fuente: Tomado de Gutierrez & de la Vara, (2008).

Con el diseño factorial completo 2^k se pueden estudiar en total los $2^k - 1$ efectos siguientes:

$$\binom{k}{1} = k, \text{ efectos principales}$$

$$\binom{k}{2} = \frac{k!}{2!(k-2)!} = \frac{k(k-1)}{2}, \text{ interacciones dobles}$$

$$\binom{k}{3} = \frac{k!}{3!(k-3)!}, \text{ interacciones triples. Y así hasta}$$

$$\binom{k}{k} = 1 \text{ interacción de los } k \text{ factores.}$$

Donde la operación $\binom{k}{r} = \frac{k!}{r!(k-r)!}$ son las combinaciones de k factores tomados de r en r (Gutierrez & de la Vara, 2008).

Estimación de contrastes, efectos y sumas de cuadrados. Cada uno de los efectos se estima a partir de su contraste, el cual a su vez se puede obtener construyendo la tabla de signos del diseño. El contraste de cada efecto se obtiene al multiplicar su columna de signos por la columna de totales expresados en la notación de Yates. Con los contrastes se procede a estimar los efectos mediante la fórmula:

$$Efecto\ ABC \dots K = \frac{1}{n2^{k-1}} [Contraste\ ABC \dots K]$$

para el cual su suma de cuadrados con un grado de libertad está dada por:

$$SC_{AB\dots K} = \frac{1}{n2^k} [Contraste\ ABC \dots K]^2$$

donde n es el número de réplicas del experimento (Gutierrez & de la Vara, 2008).

ANOVA del diseño factorial 2k. La suma de cuadrados totales (SCT) en el diseño factorial 2k se calcula como:

$$SC_T = \sum_{i=1}^{n2^k} Y_i^2 - \frac{Y_*^2}{n2^k}$$

y tiene $n2^k - 1$ grados de libertad, donde el subíndice i corre sobre el total de observaciones. La suma de cuadrados del error (SCE) se obtiene por diferencia y tiene $2k(n - 1)$ grados de libertad. Con estas dos sumas de cuadrados y las de los efectos se procede a escribir la tabla de ANOVA.

Si en la tabla de ANOVA se incluye el total de efectos que se estiman con el factorial completo 2^k , será necesario realizar cuando menos dos réplicas del experimento para estimar una suma de cuadrados del error. Sin embargo, en la mayoría de los casos sólo interesa estudiar los efectos principales y las interacciones dobles. Esto hace que cuando el número de factores es mayor o igual a cuatro ($k \geq 4$) no sea estrictamente necesario realizar réplicas.

Cabe agregar que cuando se emplea un diseño factorial 2^k , se supone que la respuesta es aproximadamente lineal en el rango de variación de cada uno de los factores estudiados. No es necesario suponer una linealidad perfecta, pero sí que no haya una curvatura muy grande. De esta manera, dado que cada factor se prueba en dos niveles, no es posible estudiar efectos de curvatura, aunque ésta exista en el proceso; para estudiar tales efectos se necesitan al menos tres niveles en cada factor. Esto no implica que de entrada sea recomendable un diseño factorial

con al menos tres niveles en cada factor, sino que en primera instancia se pueden agregar repeticiones (mínimo tres) al centro del diseño factorial 2^k , y con ellas se podrá detectar la presencia de curvatura (Gutiérrez & de la Vara, 2008).

Notación de Yates	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
(1)	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+
a	+	-	-	-	-	-	-	-	-	+	+	+	+	+	+
b	-	+	-	-	-	-	+	+	+	-	-	-	+	+	+
ab	+	+	-	-	-	+	-	-	-	-	-	-	+	+	+
c	-	-	+	-	-	+	-	+	+	-	+	+	-	-	+
ac	+	-	+	-	-	-	+	-	-	-	+	+	-	-	+
bc	-	+	+	-	-	-	-	+	+	+	-	-	-	-	+
abc	+	+	+	-	-	+	+	-	-	+	-	-	-	-	+
d	-	-	-	+	-	+	+	-	+	+	-	+	-	+	-
ad	+	-	-	+	-	-	-	+	-	+	-	+	-	+	-
bd	-	+	-	+	-	-	+	-	+	-	+	-	-	+	-
abd	+	+	-	+	-	+	-	+	-	-	+	-	-	+	-
cd	-	-	+	+	-	+	-	-	+	-	-	+	+	-	-
acd	+	-	+	+	-	-	+	+	-	-	-	+	+	-	-
bcd	-	+	+	+	-	-	-	-	+	+	+	-	+	-	-
abcd	+	+	+	+	-	+	+	+	-	+	+	-	+	-	-
e	-	-	-	-	+	+	+	+	-	+	+	-	+	-	-
ae	+	-	-	-	+	-	-	-	+	+	+	-	+	-	-
be	-	+	-	-	+	-	+	+	-	-	-	+	+	-	-
abe	+	+	-	-	+	+	-	-	+	-	-	+	+	-	-
ce	-	-	+	-	+	+	-	+	-	-	+	-	-	+	-
ace	+	-	+	-	+	-	+	-	+	-	+	-	-	+	-
bce	-	+	+	-	+	-	-	+	-	+	-	+	-	+	-
abce	+	+	+	-	+	+	+	-	+	+	-	+	-	+	-
de	-	-	-	+	+	+	+	-	-	+	-	-	-	-	+
ade	+	-	-	+	+	-	-	+	+	+	-	-	-	-	+
bde	-	+	-	+	+	-	+	-	-	-	+	+	-	-	+
abde	+	+	-	+	+	+	-	+	+	-	+	+	-	-	+
cde	-	-	+	+	+	+	-	-	-	-	-	-	+	+	+
acde	+	-	+	+	+	-	+	+	+	-	-	-	+	+	+
bcde	-	+	+	+	+	-	-	-	-	+	+	+	+	+	+
abcde	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Ilustración 6. Diseños factoriales 2^k y sus efectos de interés, $k \leq 5$.

Fuente: Tomado de Gutiérrez & de la Vara, (2008).

4.6. Diseño de experimentos como herramienta para validar modelos de simulación

Los modelos de simulación son empleados en diferentes áreas de la investigación, ya que son de gran utilidad para conocer la conducta de ciertos factores bajo diferentes contextos virtuales propiciados por el investigador. En el campo de la estadística son muy comunes los estudios de robustez, muchos de ellos utilizados para determinar el comportamiento de un estimador ante diferentes situaciones hipotéticas que pudieran presentarse en la realidad.

Los estudios experimentales de efectos fijos son utilizados para estudiar el efecto de ciertas variables controladas por el investigador, sobre una o varias variables respuesta que representen un fenómeno de interés. Esta metodología es la empleada para probar causalidad al fijar los niveles de las variables explicativas, en este caso llamadas factores, y evaluar los cambios generados sobre las variables respuestas (Montgomery, 1997).

Cada tratamiento conformado por la combinación de los diferentes niveles de los factores, genera un escenario en particular. Cuando dicho escenario corresponde a una situación virtual, propiciada por algún software especializado, se habla de los estudios de simulación (Law, 2000).

Una de las diferencias fundamentales entre este tipo de experimentos sistematizados y los experimentos físicos es que en los primeros se tiene mayor intervención sobre los factores que afectan sobre una respuesta ya que estos factores son compuestos a partir de algoritmos estocásticos predefinidos cuyos parámetros son declarados por el investigador. En contraste, los experimentos físicos generan resultados que dependen de factores que frecuentemente no es posible controlar (Salazar & Baena, 2008).

4. Metodología

El desarrollo de la metodología comprende diferentes etapas, a través de estas, se logra comprender el manejo de diseño y análisis de experimentos, estableciendo la estrategia adecuada para la manipulación de datos, obteniendo como resultado final, una metodología aplicable al campo del desarrollo del proyecto.

La metodología propuesta para esta investigación comprende las siguientes etapas:

- Identificar modelo de diseño de experimentos, aplicable al modelo de simulación basada en agentes construido.
- Proponer un diseño de experimentos aplicable al modelo de simulación basada en agentes.
- Validar el modelo de simulación basada en agentes por medio del modelo de diseño de experimentos propuesto.
- Evaluar el efecto del cambio en los niveles de los factores del diseño experimental en la variable respuesta.
- Proponer políticas que permitan disminuir el impacto del fenómeno del fraude académico en los programas de pregrado de la UNAD.

5. Modelo de simulación basada en agentes

El modelo de simulación basada en agentes fue desarrollado usando el software NetLogo, a continuación, se describe el funcionamiento del modelo.

Interfaz

En el modelo se aprecian dos secciones marcadas: La vista del entorno de simulación (Izquierda) y los controles del modelo (derecha.). En la vista se observan unas secciones de colores, cada parche de estas secciones representa un estudiante en el grupo de cada profesor (las siluetas de persona que están en el modelo).

Como controles están el número de grupos, el número de actividades por asignatura, el tamaño de la red de los estudiantes (con cuantos comparten información del curso), los niveles de percepción de fraude del profesor (laxo o estricto), la percepción de severidad de la sanción por parte de los estudiantes (Baja o alta) y la percepción de la dificultad de las pruebas por parte de los estudiantes (alta o baja). Los dos gráficos muestran el número de estudiantes aprobados y estudiantes fraudulentos por cada corrida en el tiempo.

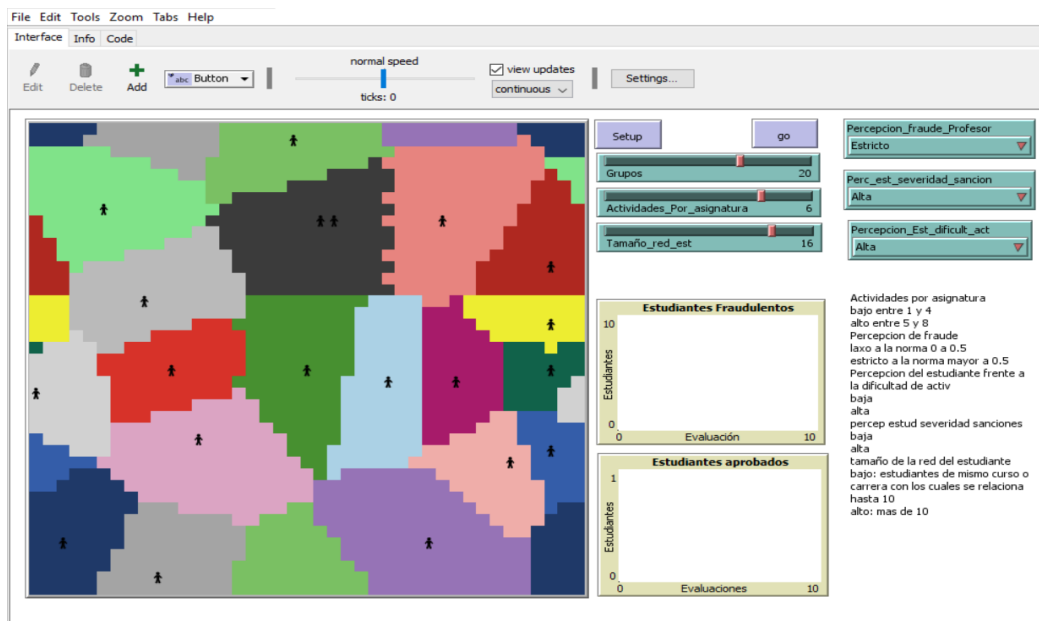


Ilustración 7. Modelo de simulación basada en agentes

Fuente: elaboración propia.

Funcionamiento del modelo

El modelo está compuesto por los agentes profesores, que interactúan con los agentes estáticos (estudiantes). Cada que hay una iteración, el estudiante que pertenecen al mismo profesor, preguntar si en promedio es más probable que hagan fraude, si hay más estudiantes que en promedio lo harán entonces aleatoriamente el estudiante decide si hacer o no fraude. Si no se supera este umbral, el estudiante no hará fraude.

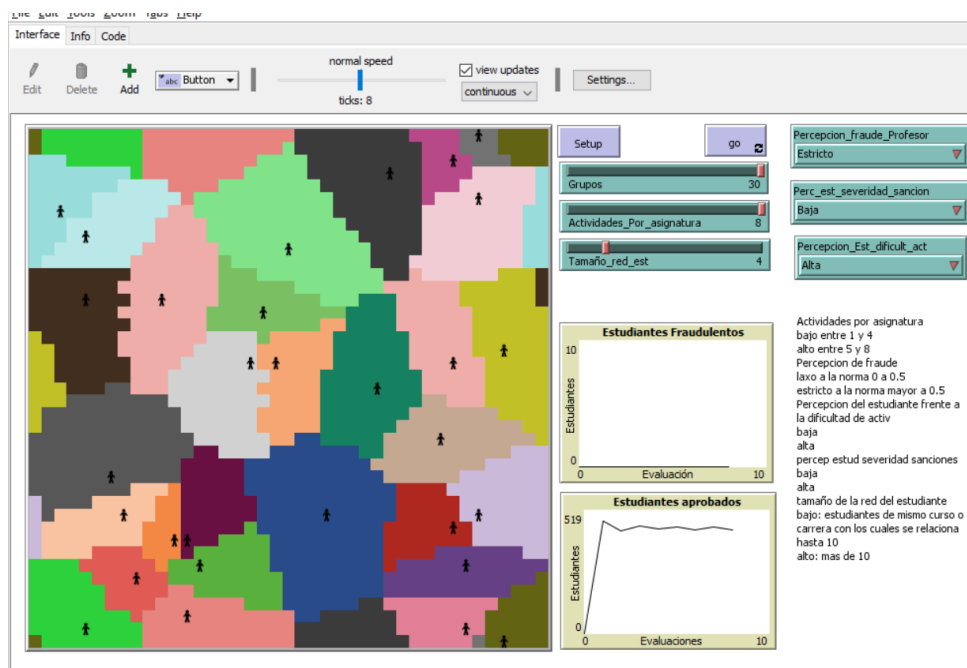


Ilustración 8. Funcionamiento del modelo de simulación basada en agentes.

Fuente: elaboración propia.

La calificación es una variable aleatoria triangular con valores (0, dificultad examen, 5) y se evalúa para todos los estudiantes en función de las características del agente profesor. El estudiante que haga fraude puede ser identificado o no, y si es identificado, en función de la percepción de fraude del profesor se procederá a penalizarlo con una nota de cero. Posteriormente se presenta la nota promedio de los estudiantes incluyendo los fraudulentos identificados y también se grafican los estudiantes fraudulentos identificados.

Presentar introducción al algoritmo (indicar si es autoría propia o sencillamente presentarlo en anexos) puede reemplazarse por el resumen de la descripción de la secuencia lógica que este realiza (que ya está en los pasos anteriores)

Código fuente del modelo:

El código fuente utilizado para el modelo de simulación en estudio se puede ver completamente en el anexo 1 de éste documento

6. Modelo de diseño de experimentos

Considero importante retomar una introducción que indique el proyecto es continuación de otro o parte de uno macro

Para analizar los resultados del modelo de simulación basada en agentes, se usará un diseño experimental del tipo Diseño Factorial 2⁵, el cual, consiste en un diseño factorial en el que se tienen una variable respuesta, cinco factores y dos niveles para cada uno de estos factores.

El primero factor es el “número de actividades por asignatura”, los niveles de este factor fueron dos; bajo (cuando la asignatura presenta cuatro o cinco actividades), y alto (cuando la asignatura presenta seis o siete actividades).

El segundo factor es “percepción de fraude del docente”, para este factor, se consideraron dos niveles; estricto (el docente se considera estricto frente a la normativa) y laxo (el docente se considera laxo frente a estas conductas).

El tercer factor es “percepción del estudiante frente a la dificultad de la actividad”, para este factor se consideraron los siguientes niveles, alta (cuando el estudiante considera que la actividad presenta un grado de dificultad alto), y baja (cuando el estudiante considera que la actividad es poco exigente o fácil).

El cuarto factor es la “percepción del estudiante frente a severidad de sanciones”, los niveles de este factor, son; alta severidad (cuando el estudiante considera que, cometer alguna conducta considerada como fraude en su actuar como estudiante de la UNAD, es castigada con severidad), y baja (cuando el estudiante considera que, cometer alguna conducta considerada como fraude en su actuar como estudiante de la UNAD, no será castigada o de ser castigado, este castigo será muy leve).

Finalmente, el quinto factor es el “tamaño de la red del estudiante” (número de compañeros con los que tiene contacto el estudiante), los niveles de este factor son; pequeña (cuando el estudiante interactúa con máximo diez estudiantes), y grande (cuando el estudiante interactúa con más de diez estudiantes).

Para estudiar los efectos del número de actividades por asignatura, la percepción de fraude del docente, la percepción del estudiante frente a la dificultad de la actividad, la percepción del estudiante frente a la severidad de las sanciones y el tamaño de la red del estudiante sobre el fenómeno del fraude académico en los programas virtuales de la UNAD, se declaró como variable respuesta el “número de estudiantes que incurrir en alguna conducta considerada como fraude”.

Finalmente, a continuación, se presenta el modelo de diseño experimental planteado:

Factores y niveles del diseño experimental propuesto	
A=	Número de actividades por asignatura (Bajo, Alto).
B=	Percepción de fraude del docente (Laxo, Estricto).
C=	Percepción del estudiante frente a la dificultad de la actividad (Baja, Alta).
D=	Percepción del estudiante frente a severidad de sanciones (Baja, Alta).
E=	Tamaño de la red del estudiante (Pequeña, Grande).

Tabla 1. Factores y niveles del diseño experimental propuesto

Fuente: elaboración propia.

Para estudiar el efecto estos cinco factores, sobre la variable respuesta y (que en este caso representa el número de estudiantes que incurren en alguna conducta considerada como fraude), se decide correr el experimento con dos corridas o réplicas. En la Tabla 2, se muestra la matriz de los tratamientos en orden aleatorio usada en el diseño experimental, por razones de simplificación, se usa muestra la matriz en unidades codificadas, en donde, por ejemplo, para el caso del factor A, codificación de sus niveles son; -1.0, para el caso del nivel “Bajo” y 1.0 para el caso del nivel “Alto”, de la misma forma, para el factor B, la codificación de sus niveles queda de la siguiente manera; -1.0, para el nivel “Laxo”, y 1.0 para el nivel “Estricto”. De la misma forma se procedió para cada uno de los demás factores y sus respectivos niveles.

Corrida	A	B	C	D	E	Y
1	1.0	-1.0	1.0	1.0	1.0	62
2	1.0	-1.0	-1.0	-1.0	1.0	49
3	1.0	1.0	-1.0	1.0	1.0	4
4	1.0	1.0	-1.0	1.0	-1.0	0
5	-1.0	1.0	1.0	1.0	1.0	0
6	1.0	1.0	1.0	1.0	1.0	0
7	-1.0	1.0	1.0	1.0	-1.0	0
8	-1.0	1.0	-1.0	1.0	1.0	0
9	-1.0	1.0	1.0	-1.0	-1.0	0
10	1.0	-1.0	-1.0	-1.0	-1.0	125
11	1.0	1.0	-1.0	-1.0	-1.0	0
12	-1.0	-1.0	-1.0	-1.0	-1.0	97
Corrida	A	B	C	D	E	Y
13	-1.0	1.0	-1.0	-1.0	1.0	0
14	-1.0	1.0	1.0	1.0	-1.0	0

15	1.0	-1.0	1.0	-1.0	1.0	136
16	-1.0	-1.0	-1.0	1.0	1.0	22
17	1.0	1.0	1.0	1.0	1.0	0
18	1.0	1.0	1.0	1.0	-1.0	0
19	1.0	-1.0	-1.0	1.0	-1.0	95
20	1.0	-1.0	1.0	1.0	1.0	131
21	1.0	-1.0	-1.0	1.0	1.0	113
22	1.0	1.0	1.0	-1.0	1.0	0
23	-1.0	-1.0	1.0	1.0	-1.0	76
24	-1.0	-1.0	1.0	-1.0	1.0	140
25	-1.0	1.0	-1.0	1.0	-1.0	7
26	1.0	1.0	-1.0	-1.0	1.0	0
27	-1.0	1.0	1.0	-1.0	-1.0	0
28	-1.0	-1.0	1.0	-1.0	1.0	78
29	1.0	1.0	-1.0	1.0	1.0	0
30	1.0	-1.0	-1.0	-1.0	1.0	103
31	-1.0	-1.0	-1.0	1.0	-1.0	94
32	-1.0	-1.0	1.0	1.0	1.0	76
33	1.0	1.0	-1.0	1.0	-1.0	0
34	-1.0	1.0	1.0	-1.0	1.0	5
35	-1.0	1.0	-1.0	-1.0	1.0	0
36	-1.0	-1.0	-1.0	-1.0	1.0	175
37	-1.0	-1.0	1.0	-1.0	-1.0	74
38	-1.0	1.0	-1.0	1.0	1.0	0
39	1.0	1.0	1.0	-1.0	1.0	0
40	-1.0	1.0	1.0	1.0	1.0	0
41	-1.0	-1.0	-1.0	1.0	1.0	104
42	1.0	1.0	1.0	-1.0	-1.0	0
43	1.0	-1.0	-1.0	1.0	1.0	110
44	-1.0	1.0	-1.0	-1.0	-1.0	0
45	1.0	1.0	-1.0	-1.0	1.0	0
46	1.0	1.0	1.0	-1.0	-1.0	11
47	-1.0	-1.0	1.0	1.0	-1.0	46
48	-1.0	1.0	-1.0	-1.0	-1.0	0
49	1.0	-1.0	1.0	-1.0	1.0	68
50	1.0	1.0	1.0	1.0	-1.0	0
51	-1.0	1.0	1.0	-1.0	1.0	0
52	-1.0	-1.0	-1.0	-1.0	-1.0	75
53	-1.0	-1.0	1.0	1.0	1.0	97
54	1.0	-1.0	1.0	-1.0	-1.0	119
Corrida	A	B	C	D	E	Y
55	-1.0	-1.0	1.0	-1.0	-1.0	67
56	-1.0	-1.0	-1.0	-1.0	1.0	95

57	1.0	-1.0	-1.0	1.0	-1.0	116
58	1.0	-1.0	1.0	1.0	-1.0	131
59	1.0	-1.0	1.0	1.0	-1.0	92
60	1.0	1.0	-1.0	-1.0	-1.0	27
61	1.0	-1.0	-1.0	-1.0	-1.0	125
62	1.0	-1.0	1.0	-1.0	-1.0	153
63	-1.0	-1.0	-1.0	1.0	-1.0	61
64	-1.0	1.0	-1.0	1.0	-1.0	0

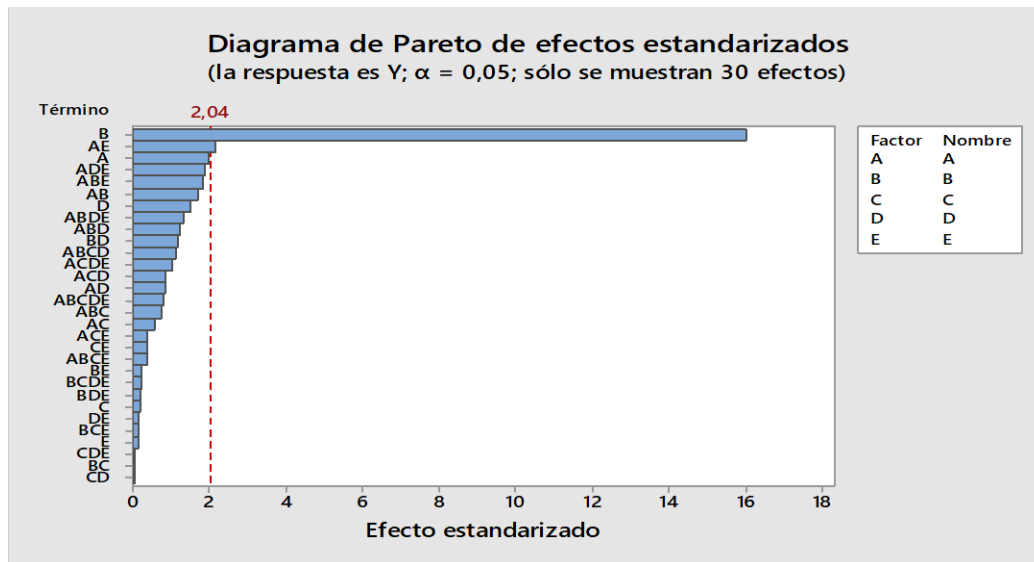
Tabla 2. Matriz de diseño en orden aleatorio (unidades codificadas).

Fuente: elaboración propia.

Finalmente, tenemos que, en total, se deben realizar 64 corridas experimentales (tendríamos 32 tratamientos y 2 réplicas por cada tratamiento), y estas corridas, se deben realizar en el orden indicado en la Tabla 2.

7. Resultados

En este capítulo, se presenta el análisis de los resultados obtenidos por el modelo de diseño de experimentos construido para analizar los resultados obtenidos por el modelo de simulación basada en agentes. Para realizar el diseño experimental, se asume un nivel de significancia de $\alpha=0.05$

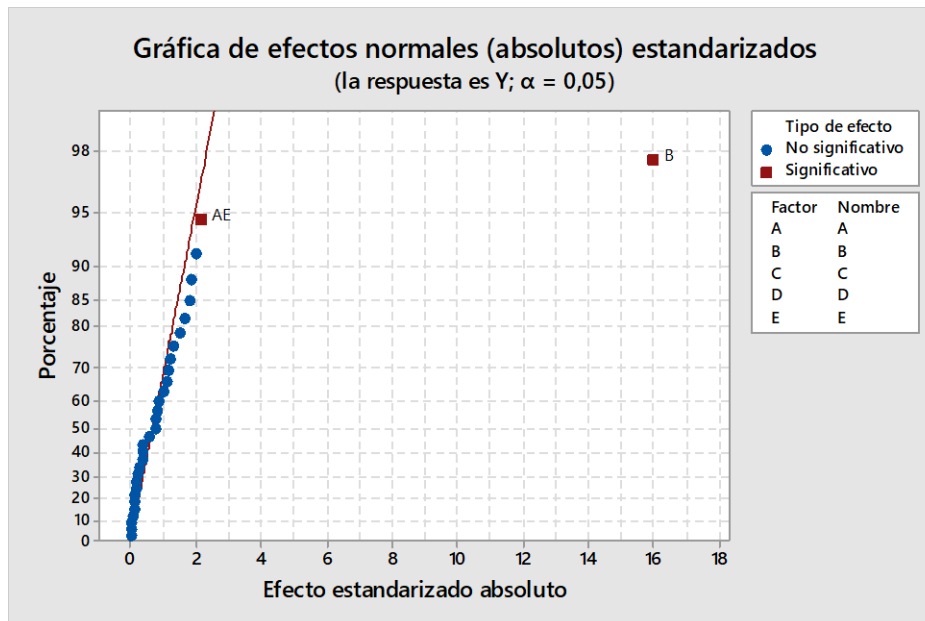


Gráfica 1. Diagrama de Pareto de efectos estandarizados

Fuente: elaboración propia.

En la gráfica 1, se muestra el diagrama de Pareto para los efectos estandarizados, recordemos que este diagrama, nos permite conocer, la magnitud y la importancia

de lo de cada uno de los efectos de los factores y sus interacciones (Gutierrez & de la Vara, 2008). En este diagrama, las barras que cruzan la línea de referencia son estadísticamente significativas. Para nuestro caso, tenemos que, las barras que representan el factor B, y la interacción AE, cruzan la línea de referencia, por lo tanto, el factor B y la interacción AE son estadísticamente significativos. Esto quiere decir, con respecto a nuestro modelo, que la percepción que tenga el estudiante sobre el docente, de si es laxo o estricto con respecto al plagio (factor B), influye directamente en el número de estudiantes que van a cometer fraude, adicionalmente, si se observa la barra correspondiente al factor B, es mucho mas larga que las de los demás factores y sus interacciones, por lo tanto, este factor, es el factor más importante de los analizados en el modelo. La interacción AE, también es estadísticamente significativa, lo cual quiere decir que, el número de actividades por asignatura (factor A) y el tamaño de la red del estudiante, ósea, el número de compañeros con los cuales tenga interacción (factor E), influyen en el número de estudiantes que terminan cometiendo fraude. Adicionalmente, la barra correspondiente al factor A, está muy cerca de la línea de referencia, por lo tanto, sería interesante realizar un análisis posterior para identificar la importancia de dicho factor.



Gráfica 2. Gráfica de efectos normales estandarizados.

Fuente: elaboración propia.

En la gráfica 2, se muestra la gráfica de efectos normales estandarizados, esta gráfica, nos permite conocer la magnitud y la importancia de los efectos de los factores y sus interacciones. En esta gráfica, los efectos que están más distantes de 0, son estadísticamente significativos. Adicionalmente, el color de los puntos y su forma, también permite identificar los efectos significativos y los que no lo son (Gutierrez & de la Vara, 2008). Para nuestro caso, igual que en el diagrama de

Pareto, los efectos significativos son los efectos del factor B y de la interacción AE. En la gráfica se pueden identificar el factor B y la interacción AE por medio de un cuadrado, mientras que los demás efectos se representan por medio de puntos, adicionalmente los cuadrados y los puntos tienen colores diferentes. Lo anterior, permite identificar claramente los factores y las interacciones significativas.

Fuente de Variabilidad	GL	Suma de cuadrados (SC)	Cuadrado Medio (MC)	Valor F	Valor p
Modelo	31	163.061,00	5.260,00	9,27	0,00000
A	1	2.268,00	2.268,00	4,00	0,05400
B	1	145.447,00	145.447,00	256,44	0,00000
C	1	19,00	19,00	0,03	0,85500
D	1	1.269,00	1.269,00	2,24	0,14400
E	1	8,00	8,00	0,01	0,90500
A*B	1	1.610,00	1.610,00	2,84	0,10200
A*C	1	179,00	179,00	0,32	0,57800
A*D	1	405,00	405,00	0,71	0,40400
A*E	1	2.665,00	2.665,00	4,70	0,03800
B*C	1	1,00	1,00	0,00	0,96300
B*D	1	763,00	763,00	1,35	0,25500
B*E	1	38,00	38,00	0,07	0,79900
C*D	1	0,00	0,00	0,00	0,97900
C*E	1	79,00	79,00	0,14	0,71200
D*E	1	10,00	10,00	0,02	0,89600
A*B*C	1	320,00	320,00	0,56	0,45800
A*B*D	1	848,00	848,00	1,50	0,23000
A*B*E	1	1.903,00	1.903,00	3,36	0,07600
A*C*D	1	425,00	425,00	0,75	0,39300
A*C*E	1	79,00	79,00	0,14	0,71200
A*D*E	1	1.969,00	1.969,00	3,47	0,07200
B*C*D	1	0,00	0,00	0,00	0,97900
B*C*E	1	8,00	8,00	0,01	0,90500
B*D*E	1	19,00	19,00	0,03	0,85500
C*D*E	1	2,00	2,00	0,00	0,95400
A*B*C*D	1	709,00	709,00	1,25	0,27200
A*B*C*E	1	79,00	79,00	0,14	0,71200
A*B*D*E	1	953,00	953,00	1,68	0,20400
A*C*D*E	1	594,00	594,00	1,05	0,31400
B*C*D*E	1	35,00	35,00	0,06	0,80700

Fuente de Variabilidad	GL	Suma de cuadrados (SC)	Cuadrado Medio (MC)	Valor F	Valor p
A*B*C*D*E	1	356,00	356,00	0,63	0,43400
Total Error	32	18.149,00	567,00		
Total	63	181.211,00			

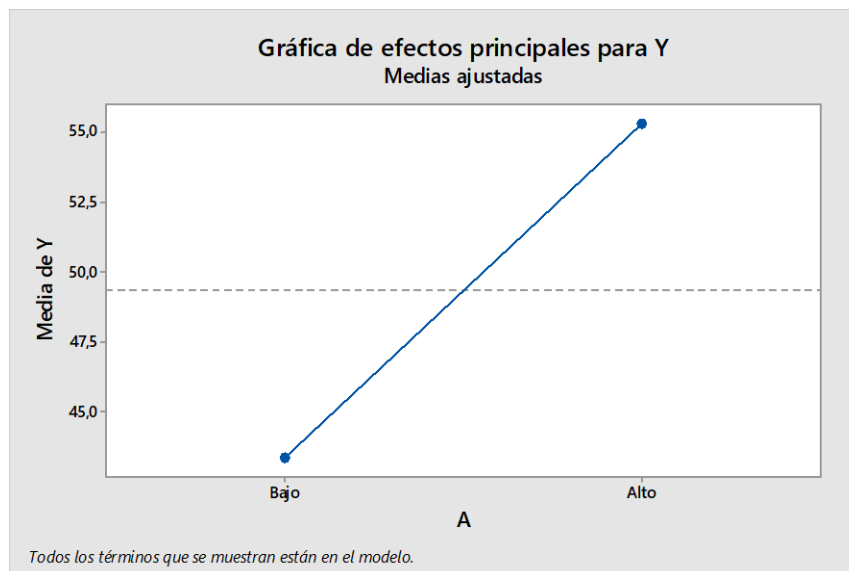
Tabla 3. ANOVA para el modelo de fraude académico.

Fuente: elaboración propia.

En la Tabla 3, se muestra la tabla ANOVA (análisis de varianza) del modelo de fraude académico desarrollado, para el análisis de los resultados, se eligió el Valor-p como estadístico de prueba, y como criterio de decisión, si el Valor-p $> \alpha=0,05$ el efecto de los factores, las interacciones y el modelo, son significativos. Teniendo en cuenta lo anterior, el modelo desarrollado es estadísticamente significativo, el efecto del factor B y de la interacción AE, son igualmente significativos. Mientras que, para el factor A, el Valor-p=0,05400 lo cual indica que, está muy cercano al valor de $\alpha = 0.05$, ósea, está muy cerca al criterio de rechazo. Por lo tanto, valdría la pena hacer un análisis mucho más detenido sobre la significancia del efecto de dicho factor.

Análisis de los efectos principales:

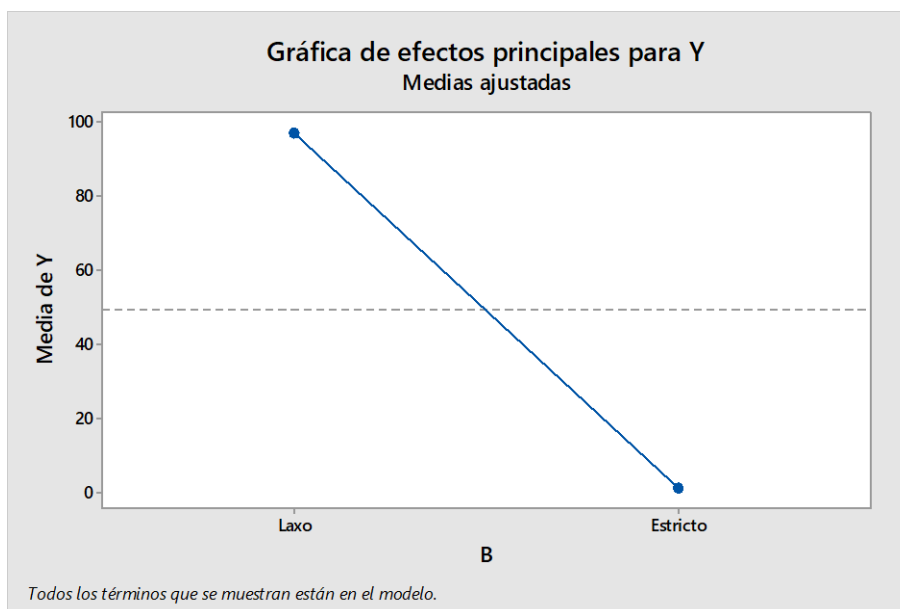
Para realizar el análisis de los efectos principales, se usa una gráfica de efectos principales, en esta gráfica, se grafican las medias para cada nivel de un factor. Esta gráfica se interpreta de la siguiente manera; si la línea que conecta los puntos de un factor, es horizontal, ósea, paralela al eje X, entonces no hay un efecto principal y cada nivel del factor, tiene el mismo efecto sobre la variable respuesta, no se observa un cambio en la variable respuesta cuando pasamos de un nivel a otro dentro de un factor. Sin embargo, si esta línea no es horizontal, esto indica que, existe un efecto principal y cada nivel del factor, afecta de forma diferente a la variable respuesta. Entre más inclinada sea la pendiente de la línea, el efecto del factor sobre la variable respuesta será mayor (Gutierrez & de la Vara, 2008).



Gráfica 3. Efectos principales para Y: factor A.

Fuente: elaboración propia.

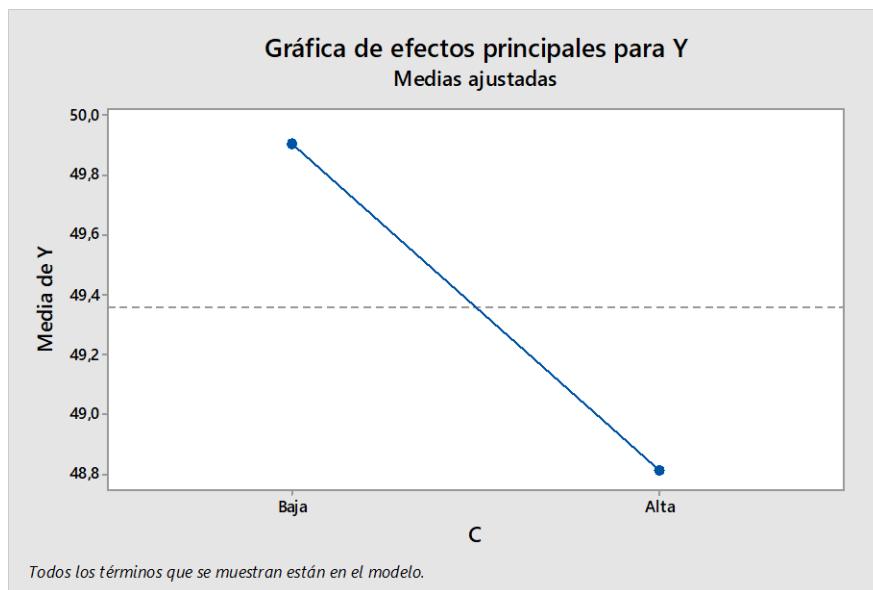
En la gráfica 3, se muestra el efecto del factor A (que representa el número de actividades por curso en un período académico), sobre la variable respuesta Y (que representa el número de estudiantes que cometen fraude por periodo académico), en esta gráfica, se puede ver que, cuando el número de actividades se incrementa, el número de estudiantes que cometen fraude también se incrementa. Dicho de otro modo, si un estudiante en sus asignaturas tiene muchas actividades, es más probable que recurra a conductas consideradas fraudulentas en el desarrollo de sus actividades.



Gráfica 4. Efectos principales para Y: factor B.

Fuente: elaboración propia.

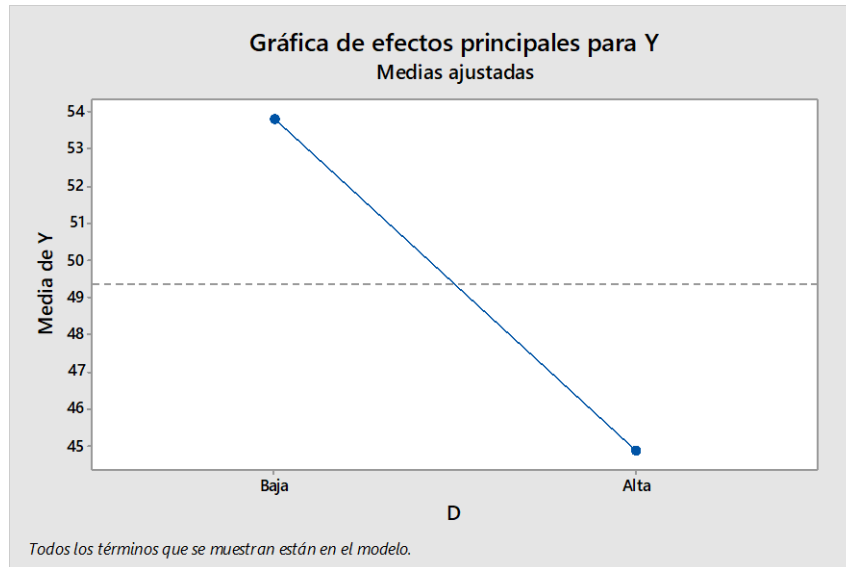
En la gráfica 4, se muestra el efecto del factor B (percepción que tiene el estudiante sobre la severidad del docente frente a conductas consideradas fraudulentas en el desarrollo de las actividades), sobre la variable respuesta Y (que representa el número de estudiantes de cometen fraude por periodo académico), en esta gráfica, se puede ver que, si el estudiante considera que el docente es laxo, ósea, no castiga las conductas fraudulentas o el castigo es leve, el número de estudiantes que comete fraude es alto, mientras que, cuando el estudiante considera que el docente es estricto ante una conducta fraudulenta en las actividades, el número de casos es cero, lo cual indica que, para evitar conductas fraudulentas en las actividades, nuestros docentes deben castigar cada caso identificado y de forma severa, esto haría que los estudiantes piensen antes de incurrir en una conducta como esta.



Gráfica 5. Efectos principales para Y: factor C.

Fuente: elaboración propia.

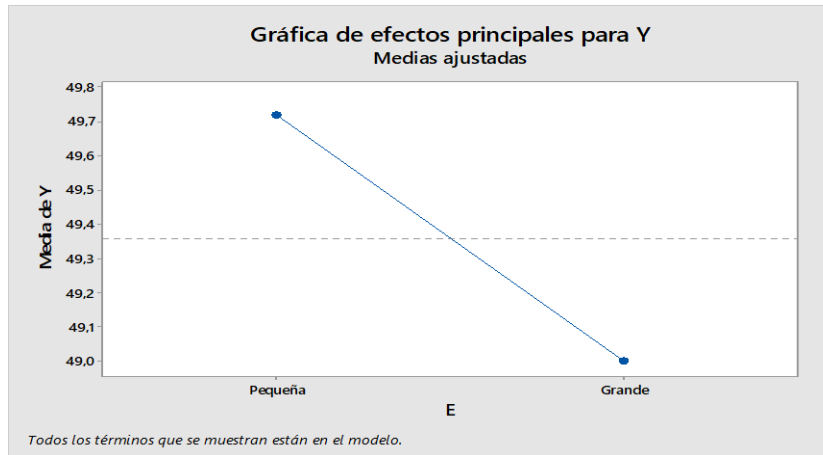
En la gráfica 5 se muestra el efecto del factor C (dificultad de la actividad), sobre la variable respuesta Y (que representa el número de estudiantes de cometen fraude por periodo académico), en esta gráfica, se puede ver que, aunque cuando se pasa del nivel bajo al alto (recordemos que el nivel bajo, representa cuando una actividad se considera fácil, mientras que el nivel alto, cuando una actividad se considera difícil por el estudiante), se presenta una disminución en el número de estudiantes que cometen fraude, las medias son muy parecidas. En el nivel Bajo del factor, la media es de 49.8, mientras que en el nivel alto es de 48.8, lo cual indica que, estadísticamente el efecto del factor no es significativo como se había demostrado en los análisis anteriores.



Gráfica 6. Efectos principales para Y: factor D.

Fuente: elaboración propia.

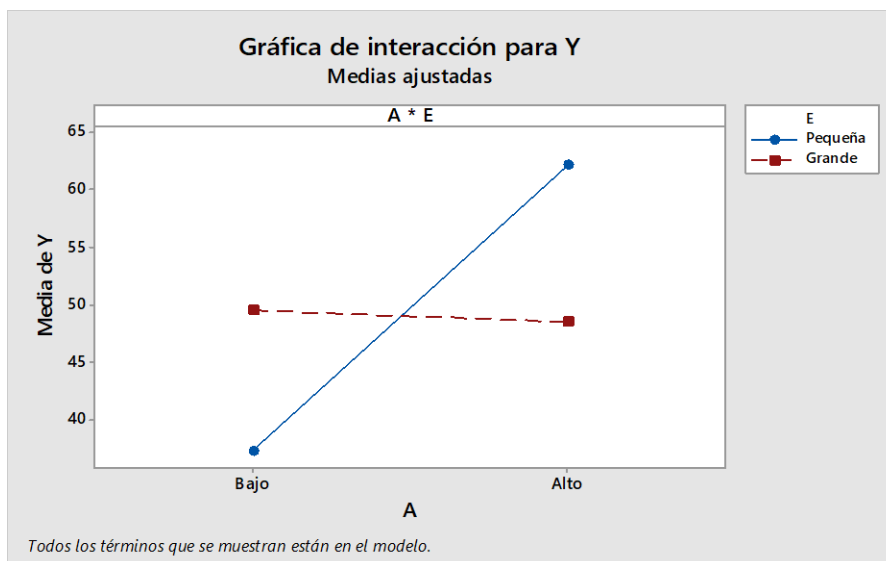
En la gráfica 6 se muestra el efecto del factor D (severidad de las sanciones), sobre la variable respuesta Y (que representa el número de estudiantes que cometen fraude por periodo académico), en esta gráfica, se puede ver que, si el estudiante considera que, al cometer fraude, la institución no lo va a castigar o el castigo va a ser leve, el estudiante será propenso a incurrir en conductas consideradas fraudulentas. Mientras que, si el estudiante considera que la institución castiga de forma severa este tipo de conductas, el número de estudiantes que incurrirán en ellas, va a disminuir, sin embargo, el problema no se va a acabar, dado que, la media es alta (es de 45), lo cual indica que, aunque este factor es importante, no es motivo de persuasión para todos los estudiantes. Esto, se podría explicar porque, así la norma sea estricta, al final, quienes identifican en primera instancia las conductas fraudulentas son los docentes y si estos no aplican las sanciones, los estudiantes seguirán incurriendo en ellas.



Gráfica 7. Efectos principales para Y: factor E.

Fuente: elaboración propia.

En la gráfica 7 se muestra el efecto del factor E (cantidad de estudiantes con los que tiene interacción un estudiante durante el período académico), sobre la variable respuesta Y (que representa el número de estudiantes que cometen fraude por periodo académico), en esta gráfica, se puede ver que, aunque existe un efecto, el efecto es despreciable, dado que, las medias de la respuesta (número promedio de estudiantes que cometen fraude), cuando se pasa de un nivel a otro del factor son muy parecidas, por ejemplo, cuando el estudiante interactúa con pocas personas (nivel “Pequeña” del factor C), el número promedio de estudiantes que comete fraude es de 49.7, mientras que cuando el estudiante interactúa con muchos compañeros (el nivel del factor es “Grande”), el promedio de estudiantes que cometen fraude es de 49.



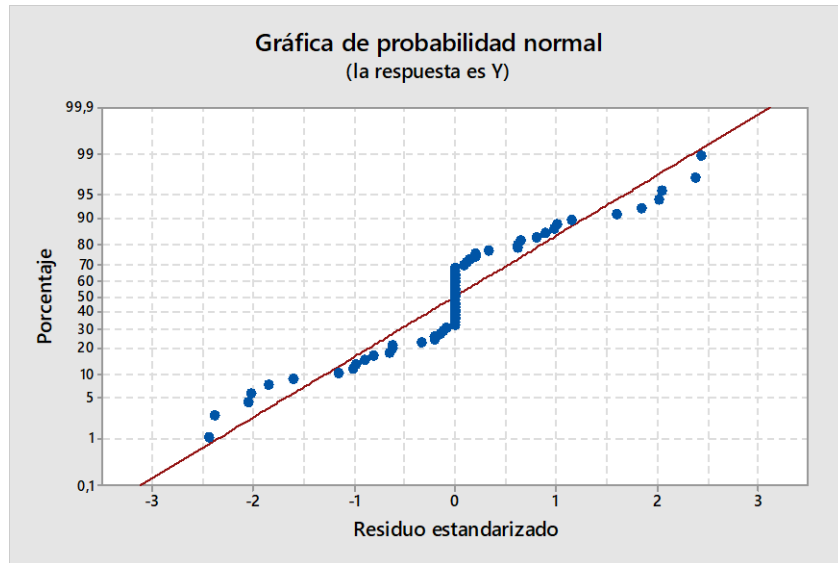
Gráfica 8. Efecto interacción para Y: interacción AE.

Fuente: elaboración propia.

En la gráfica 8 se muestra el efecto de la interacción de los factores AE sobre la variable respuesta Y (que representa el número de estudiantes de cometen fraude por periodo académico), en esta gráfica, se puede ver que, si E es “Pequeña”, ósea, el número de estudiantes con los que un estudiante interactúa durante el período académico es pequeño (menor a 10), y el número de actividades por asignatura aumenta de “Bajo” a “Alto”, el número de estudiantes que incurrir en conductas fraudulentas, también aumentaría. Por otro lado, si la red del estudiante es grande (el estudiante tiene contacto con más de 10 estudiantes durante el período, el número de actividades, ósea, el factor A, sería indiferente, dado que, se obtendría el mismo valor en la variable respuesta.

Validación de los supuestos del modelo:

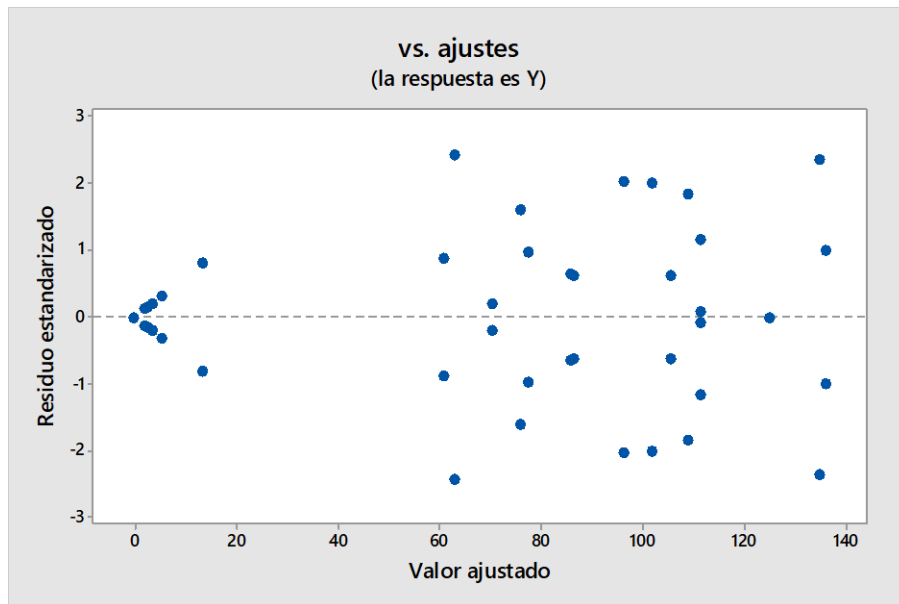
Para la validación de los supuestos del modelo de diseño de experimentos, se usan los siguientes gráficos; el gráfico de probabilidad normal de los residuos, gráfico de los residuos vs predichos, y el gráfico de residuos vs orden de las corridas. El gráfico de probabilidad normal de los residuos, se usa para comprobar el supuesto de que los residuos están distribuidos normalmente, esta gráfica debe seguir aproximadamente una línea recta. El gráfico de residuos vs predichos, se usa para comprobar el supuesto de que los residuos están distribuidos aleatoriamente y tienen una varianza constante, para que se cumpla el supuesto, en esta gráfica, todos los puntos deben caer aleatoriamente en ambos lados de 0, sin patrones identificables. Finalmente, el gráfico de residuos vs orden, permiten comprobar el supuesto de que los residuos son independientes unos de otros, para que esto ocurra, los residuos no deben mostrar tendencias o patrones cuando se muestran en orden de tiempo, si se presentan patrones, esto puede indicar que, los residuos cercanos entre sí podrían estar correlacionados y, por lo tanto, no serían independientes (Gutierrez & de la Vara, 2008).



Gráfica 9. Gráfico de probabilidad normal de los residuos estandarizados.

Fuente: elaboración propia.

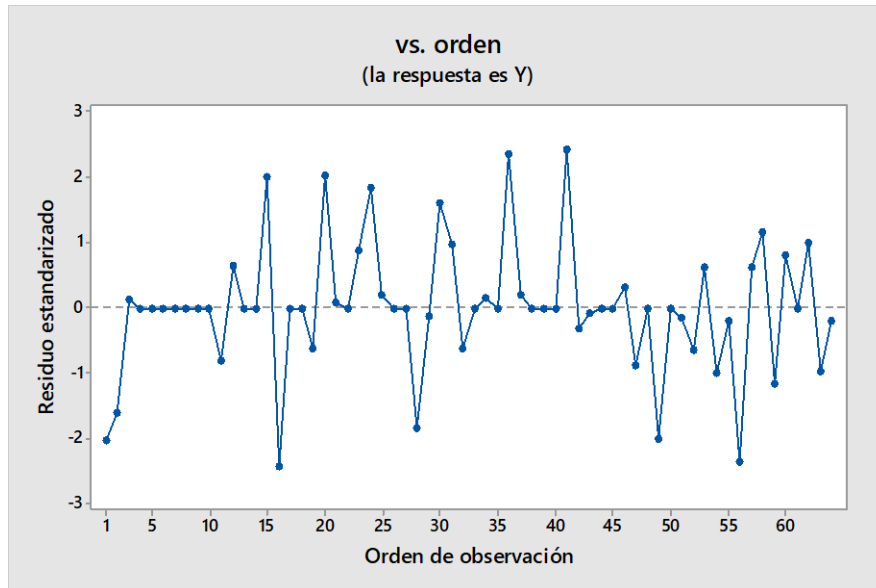
En la gráfica 9, se muestra que los residuos se ajustan a una línea recta, por lo tanto, se cumple el supuesto de normalidad de los residuos.



Gráfica 10. Gráfico de los residuos vs predichos

Fuente: elaboración propia.

En la gráfica 10, se muestra que, el supuesto de que los residuos están distribuidos aleatoriamente y tienen una varianza constante, se cumple. Dado que, todos los puntos están distribuidos a un lado y otro del 0, y no se evidencia un patrón en los datos.



Gráfica 11. Gráfico de los residuos estandarizados vs orden.

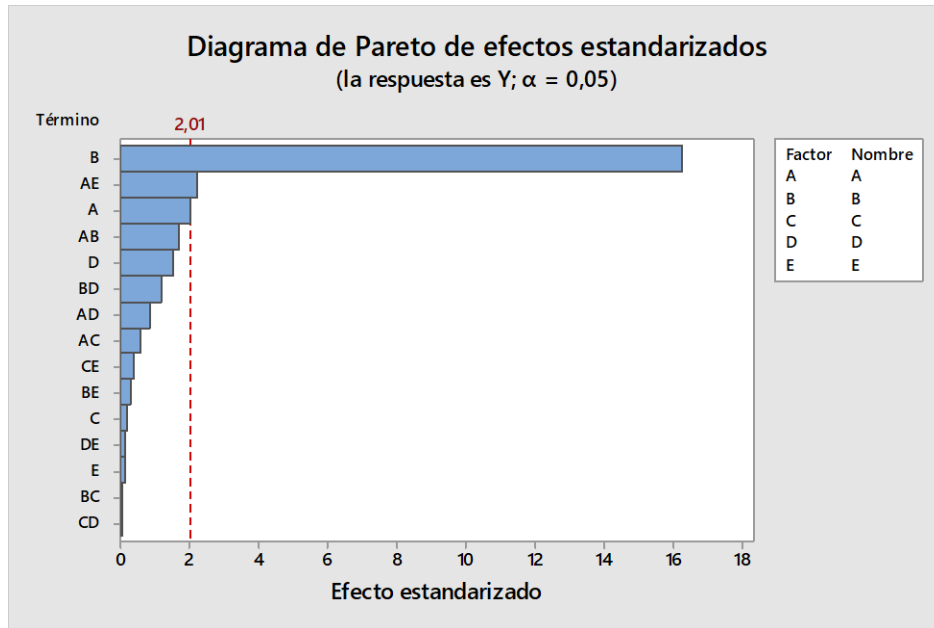
Fuente: elaboración propia.

En la gráfica 11, se muestra que, los residuos se distribuyen sin ningún patrón identificable, a un lado y otro de la línea central, por lo tanto, se cumple el supuesto de que los residuos son independientes unos de otros.

Teniendo en cuenta el análisis anterior, tenemos que, se cumplen los supuestos del modelo de diseño de experimentos.

Modelo ajustado

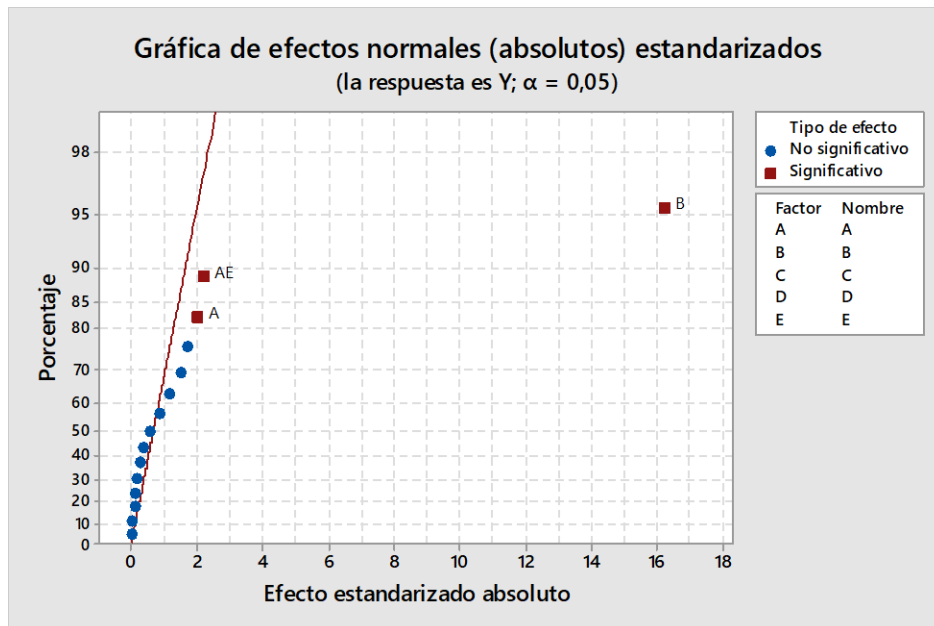
Para mejorar la precisión del modelo, se ajustó el modelo eliminando las interacciones que resultaron no significativas en el modelo original. A continuación, se presentan los resultados obtenidos.



Gráfica 12. Diagrama de Pareto de efectos estandarizados modelo ajustado

Fuente: elaboración propia.

En la gráfica 12, se muestra el diagrama de Pareto para los efectos estandarizados para el modelo ajustado. Al igual que en el diagrama del modelo original se observa que el efecto del factor B y de la interacción AE son significativos, con respecto al factor A, no es claro si su efecto sobre la variable respuesta es estadísticamente significativo.



Gráfica 13. Gráfica de efectos normales estandarizados del modelo ajustado

Fuente: elaboración propia.

En la gráfica 13, se muestra la gráfica de efectos normales estandarizados del modelo ajustado. En este caso, los efectos significativos son los efectos del factor A, del factor B y de la interacción AE. En este diagrama se muestra que, a diferencia del modelo original, el efecto del factor A es estadísticamente significativo.

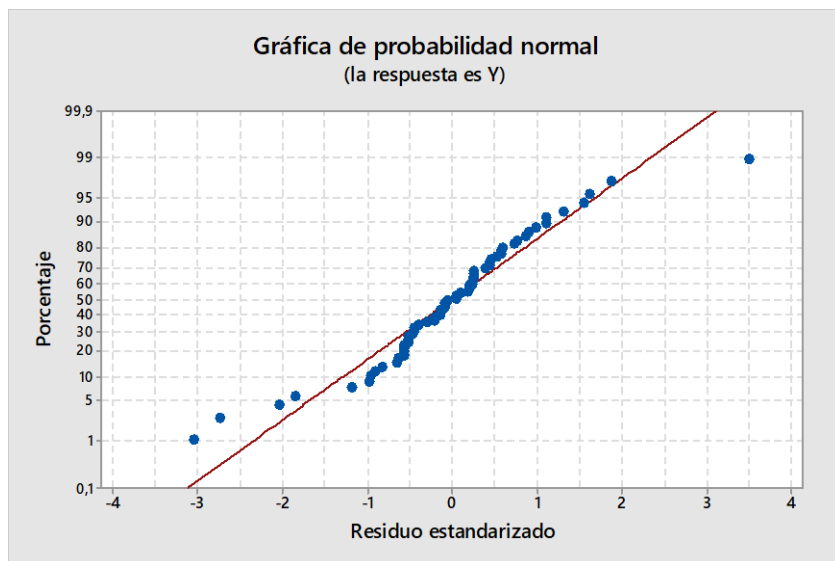
Fuente de Variabilidad	GL	Suma de cuadrados (SC)	Cuadrado Medio (MC)	Valor F	Valor p
Modelo	6	151.677,00	25.279,00	48,79	0,0000
A	1	2.268,00	2.268,00	4,38	0,0410
B	1	145.447,00	145.447,00	280,71	0,0000
C	1	19,00	19,00	0,04	0,8480
D	1	1.269,00	1.269,00	2,45	0,1230
E	1	8,00	8,00	0,02	0,9000
A*E	1	2.665,00	2.665,00	5,14	0,0270
Error	57	29.534,00	518,00		
Falta de ajuste	25	11.385,00	455,00	0,80	0,7120
Error puro	32	18.150,00	567,00		
Total	63	181.211,00			

Tabla 4. ANOVA para el modelo ajustado.

Fuente: elaboración propia.

En la Tabla 4, se muestra la tabla ANOVA del modelo ajustado. Teniendo en cuenta los resultados de este modelo, tenemos que, el modelo desarrollado es estadísticamente significativo, el efecto del factor A, del factor B y de la interacción AE, son igualmente significativos.

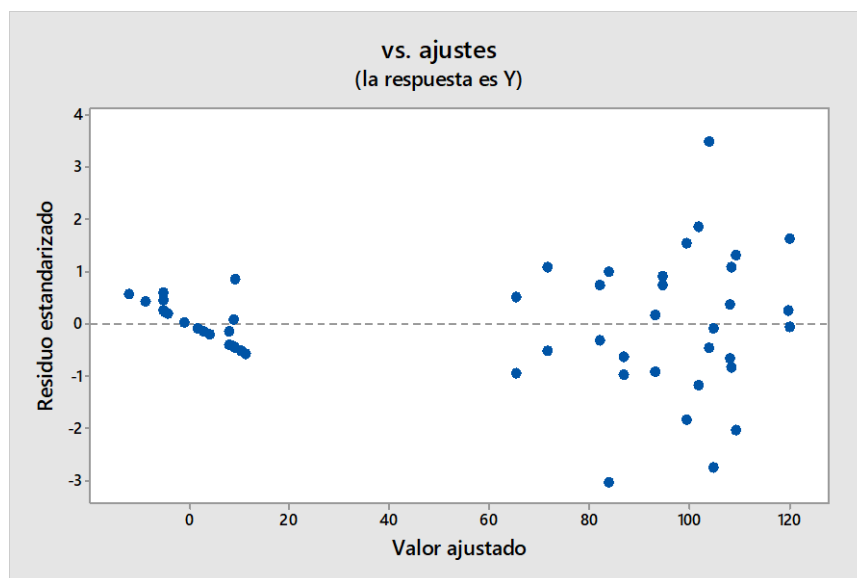
Supuestos del modelo ajustado:



Gráfica 14. Gráfico de probabilidad normal de los residuos estandarizados del modelo ajustado.

Fuente: elaboración propia

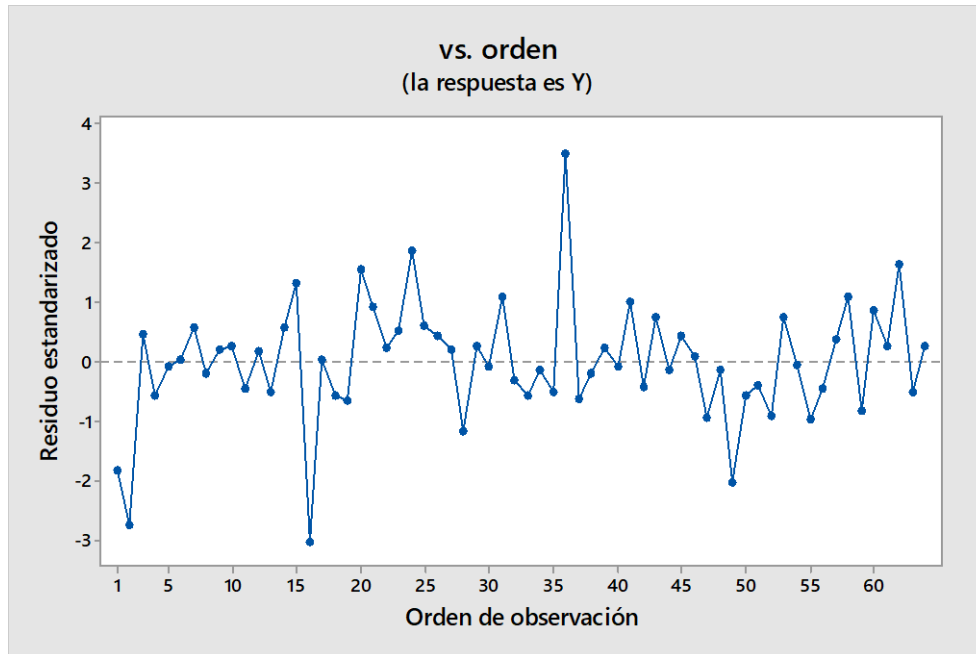
En la gráfica 14, se muestra que los residuos del modelo ajustado se ajustan a una línea recta, y aunque existen tres valores separados de los demás datos, se puede concluir que el supuesto de normalidad de los residuos se cumple.



Gráfica 15. Gráfico de los residuos vs predichos del modelo ajustado.

Fuente: elaboración propia.

En la gráfica 15, se muestra que, el supuesto de que los residuos están distribuidos aleatoriamente y tienen una varianza constante, se cumple para el modelo ajustado. Dado que, todos los puntos están distribuidos a un lado y otro del 0, y no se evidencia un patrón en los datos.



Gráfica 16. Gráfico de los residuos estandarizados vs orden del modelo ajustado

Fuente: elaboración propia.

En la gráfica 16 se muestra que, los residuos del modelo ajustado se distribuyen sin ningún patrón identificable, a un lado y otro de la línea central, por lo tanto, se cumple el supuesto de que los residuos son independientes unos de otros.

Teniendo en cuenta el análisis anterior, tenemos que, se cumplen los supuestos del modelo de diseño de experimentos ajustado.

Finalmente, para predecir el comportamiento esperado de la variable respuesta o para evaluar el efecto de cualquier combinación que se quiera sobre la variable respuesta, se presenta el modelo de regresión ajustado:

$$Y = 49.36 + 5.95X_1 - 47.67X_2 - 0.55X_3 - 4.45X_4 - 0.36X_5 - 6.45X_1X_5$$

Donde, X_1 es el factor A, X_2 es el factor B, X_3 es el factor C, X_4 es el factor D, X_5 es el factor E.

8. Conclusiones

En esta investigación, se realizó la construcción de un modelo de diseño de experimentos a partir de un modelo de simulación, para analizar el fenómeno del fraude académico en los programas virtuales de la Universidad Nacional Abierta y a Distancia. Este primer acercamiento al fenómeno, le permitirá a la UNAD y a aquellas Universidades que ofrecen programas virtuales, identificar algunas de las variables más importantes que afectan el número de casos de plagio que se puedan presentar, y con esto, plantear estrategias que permitan disminuir y/o erradicar el número de casos de fraude que se presentan cada semestre en dichas instituciones.

A partir de los resultados obtenidos por el modelo de diseño de experimentos, tenemos que, el factor más importante a la hora de determinar el número de estudiantes que cometen fraude es la percepción que tenga el estudiante sobre qué tan laxo o que tan estricto frente a las conductas tipificadas como fraude son los docentes. En este sentido, si el docente es estricto frente a una conducta considerada como fraude, por ejemplo, copiar en un examen, los estudiantes son más temerosos de cometer este tipo de acciones. Sin embargo, si los estudiantes consideran que el docente es laxo frente a este tipo de faltas, va a aumentar el número de estudiantes que cometan fraude en su vida académica. Lo anterior, se entiende, si tenemos en cuenta que los estudiantes cometen este tipo de conductas, no por falta normas o sanciones, porque estas existen en la UNAD y en todas las instituciones de educación superior, sino porque tan estricto sea el docente en caso de presenciar estas conductas.

De los resultados del modelo, también tenemos que, el número de actividades que se le asignen durante un período académico a los estudiantes en un curso, y la cantidad de estudiantes con los que un estudiante tenga contacto en ese período académico, va a determinar si el estudiante comete fraude en sus actividades académicas o no. Adicionalmente, el factor “número de actividades por asignatura”, a partir de un análisis más profundo, resultó ser significativo, lo cual se explica en el hecho de, si un curso tiene muchas actividades, y el estudiante no tiene tiempo, es propenso a conductas como, mandar a hacer un trabajo, copiarle el trabajo a un compañero, conseguir a alguien para que le ayude a responder los exámenes, entre otros.

Mediante el desarrollo de este proyecto investigativos donde se obtuvieron resultados significativos y de gran utilidad para el análisis interno de la Universidad Nacional Abierta y a Distancia UNAD en cuanto al fraude académico, se logra poner en práctica la simulación basada en agentes por medio de la técnica de diseño de experimentos ²⁵ para éste caso, donde por medio del software NetLogo se llega a la respuesta de la pregunta clave del planteamiento del problema, puesto que en los resultados obtenidos se evidencia de una manera clara cómo es la dinámica del

fraude académico en los estudiantes de los diferentes programas académicos de la universidad frente a la normatividad vigente por parte de la misma.

9. Recomendaciones

El modelo de simulación a partir del cual se construye el diseño experimental fue construido a partir de los datos obtenidos en una prueba piloto a base de encuestas aplicadas a un grupo de estudiantes de la Escuela de Ciencias Básicas, Tecnología e Ingeniería (ECBTI) de la Universidad Nacional Abierta y a Distancia (UNAD). Esta prueba piloto, sirve como un primer acercamiento al problema, sin embargo, se recomienda en estudios posteriores, realizar un estudio mucho más profundo en donde se tenga en cuenta, no sólo la opinión de los estudiantes, sino también la opinión de los docentes. Esto, permitiría validar las conclusiones obtenidas hasta ahora y conocer aún más el problema y plantear propuestas de solución que permitan disminuir su incidencia e impacto sobre la comunidad académica de los programas virtuales.

Una recomendación importante a tener en cuenta en posteriores estudios investigativos con base a este proyecto es involucrar más actores de la universidad que tengan estrecho relacionamiento con los estudiantes, como también la influencia de otros factores no tenidos en cuenta y que posiblemente pueden afectar significativamente el comportamiento del fenómeno en estudio y así mismo llegan a obtener más resultados importantes para el análisis y accionamiento de nuevas estrategias para contrarrestar los impactos que genera este fenómeno académico..

10. Referencias

AGUILAR PASSANO, María del Pilar. Plagio, un problema emergente en la escritura académica. *Revista Uruguaya de Cardiología*: (2019). 139-140.

BOYTHA, G. El plagio según la OMPI. *Revista de propiedad intelectual*, (1980). 21-62.

COOS BU, Raúl. *Simulación: un enfoque práctico*. México: LIMUSA, S.A. de C.V. 2003.

GARCÍA DUNNA, Eduardo. GARCÍA REYES, Heriberto. CÁRDENAS BARRÓN, Leopoldo Eduardo. *Simulación y análisis de sistemas con ProModel*. México: Pearson Educación de México S.A. 2006.

GARUD SUHAS, Sushant, I. A, Karimi. & KRAFT, Markus. Smart Sampling Algorithm for Surrogate Model Development. *Computers & Chemical Engineering*, 2017. 104-114.

GUTIERREZ PULIDO, Humberto. & DE LA VARA SALAZAR, Román. *Análisis y diseño de experimentos*. México, D.F. 2008.

IZQUIERDO, Luis. GALÁN ORDAX, José Manuel. SANTOS, José. & DEL OLMO MARTINEZ, Ricardo. Modelado de sistemas complejos mediante simulación basada en agentes y mediante dinámica de. *Revista de Metodología de las Ciencias Sociales*, (2008). 85-112.

GARCÍA VALDECASAS MEDINA, José Ignacio. La simulación basada en agentes: una nueva forma de explorar los fenómenos sociales. *Revista Española de Investigaciones Sociológicas*, (2011). 91-110.

AVERILL, Law. *Simulation Modeling and Analysis*. New York. 2000.

KUMAR, Kumar. WHITTAKER, Andrew. Cross-platform implementation, verification and validation of advanced mathematical models of elastomeric seismic isolation bearings. *Engineering Structures*, (2018). 1-18.

ARROYO MENÉNDEZ, Millán & HASSAN COLLADO, Samer. Simulación de procesos sociales basada en agentes software. *EMPIRIA: Revista de Metodología de Ciencias Sociales*. (2007). Vol. 14, p139-161. 23p.

PIERA EROLES, Miquel Àngel. *Cómo mejorar la logística de su empresa mediante la simulación*. Madrid: Díaz de Santos S.A. 2006.

MONTGOMERY, Douglas. *Design and Analysis of Experiments*. New York.1997.

ROJAS CHAVARRO, Miguel Ángel. Plagio en el ambito académico. *Rev. Col. Anest.* (2010). 537-538.

SALAZAR, Juan Carlos. & BAENA ZAPATA, Armando. *Análisis Y Diseño De Experimentos Aplicados A Estudios de Simulación*. Medellín. 2008.

SANABRIA, Luis Eduardo. Conceptualización Jurídica del Plagio en Colombia. *Revista Colombiana de Cirugía*. (2014). 88-97.

SARGENT, Robert. *Verification And Validation Of Simulation Models*. USA. 2010.

SARGENT, Robert. Verification And Validation Of Simulation Models. *Proceedings of the 2011 Winter Simulation Conference*. Syracuse, NY. (2011). (pág. 183).

S, Schlesinger. *Terminology for Model Credibility*. 1979.

SOTO RODRÍGUEZ, Armando. El plagio y su impacto a nivel académico y profesional. *Revista e-Ciencias de la Información*, (2012). 1-15.

ECHAVEGUREN, Tomás. Conceptos para la modelación de sistemas de gestión de activos viales usando simulación basada en agentes. *Revista Ingeniería de Construcción*. (2016). 47-56.

ANEXOS

Anexo 1: Código fuente del modelo:

::ask - hacer un ciclo sobre un conjunto de elementos con alguna característica

:: SHOW muestra el valor de una variable

:: Prob Fraude: Probabilidad de incurrir fraude

:: Perc_severidad: Percepcion de severidad

globals[

available-colors

profe-list

]

patches-own [

tipo ;; Tipo de agente: Estudiante=0, Tutor=1 (aplicacion de la norma)

:: Estudiantes

Prob_fraude ;; Prob Fraude: Probabilidad de incurrir fraude

Perc_severidad ;; Perc_severidad: Percepcion de severidad

Perc_dificul_exam ;; Percepcion de la dificultad de las tareas

tutor-related ;; Guarda el profesor al cual debe entregar info

own_color ;; el color del parche

Nota_est ;; calificacion del estudiante

Tamaño_red ;; cuantos estudiantes conoce

]

turtles-own[

:: Tutor

Perc_gravedad_fraude ;; Percepción de gravedad de fraude por parte del tutor

Frecuencia_tareas ;; Cada cuando se hace un examen

Num_estudiantes ;; total estudiantes asociados

Prop_Est_fraude;; del total del grupo cuantos hacen fraude

]

```
::INICIALIZAR
```

```
to Setup
```

```
  ca
```

```
  reset-ticks
```

```
  set-default-shape turtles "person"
```

```
  make-world
```

```
end
```

```
::DESCRIPCION DEL PROCESO
```

```
to go
```

```
  ;;COMPORTAMIENTO
```

```
  ask patches with [tipo = 0] [
```

```
    ;; Si el parche tiene vecinos que pertenecen al mismo profesor, preguntar si en promedio es mas probable que hagan
```

```
    ;; fraude, si si lo son, entonces un random con limite la prob del parche
```

```
    let tutor [who] of min-one-of turtles [distance myself]
```

```
    ifelse count patches with [own_color = pcolor and tutor-related = tutor] >
```

```
    Tamaño_red_est
```

```
    [ set Tamaño_red (patch-set n-of Tamaño_red_est patches with [own_color = pcolor and tutor-related = tutor] )]
```

```
    [ set Tamaño_red (patch-set n-of count patches with [own_color = pcolor and tutor-related = tutor] patches with [own_color = pcolor and tutor-related = tutor] )]
```

```
    if mean [Prob_fraude] of Tamaño_red > Prob_fraude
```

```
    [
```

```
      if random-float 1 < Prob_fraude
```

```
      [
```

```
        set pcolor pcolor + 1
```

```
        ; hace fraude??
```

```
        set Prob_fraude 0
```

```
      ]
```

```

]
]
;; se evalua cuantos estudiantes hacen fraude identificado y se les pone
calificacion
ask turtles [
  ;; se identifica el tutor
  let tutor [who] of min-one-of turtles [distance myself]
  ;show tutor
  ;; se calcula la proporcion de estudiantes que han hecho fraude
  set Prop_Est_fraude 1 - count patches with [own_color = pcolor and tutor-
related = tutor] / [Num_estudiantes] of turtle tutor
  set Prop_Est_fraude Prop_Est_fraude + ([Frecuencia_tareas] of turtle tutor) /
100
  ; show count patches with [own_color = pcolor and tutor-related = tutor]
  ; show [Num_estudiantes] of turtle tutor
  ;; si la Prop_Est_fraude > Perc_gravedad_fraude todos pierden, de lo contrario
nada
  ifelse Prop_Est_fraude > Perc_gravedad_fraude[
    ask patches with [tutor-related = tutor and own_color != pcolor][
      set pcolor pcolor - 2
      set Nota_est -1]
    ask patches with [tutor-related = tutor and own_color = pcolor][
      ;set pcolor own_color
      set Nota_est triang 0 Perc_dificul_exam 5]
  ]
[
ask patches with [tutor-related = tutor ][
  ;set pcolor own_color
  set Nota_est triang 0 Perc_dificul_exam 5
]
]

```

```

]
]
  ask patches [;with [tipo != 2] [
    set pcolor own_color
  ]
if ticks > 50
[
  stop
]
tick ;; para que grafique
end
;; Construccion del mundo
to make-world
  set available-colors shuffle filter [ ?1 -> (?1 mod 10 >= 2) and (?1 mod 10 <= 8) ]
    n-values 160 [ ?1 -> ?1 ]
  set profe-list n-values 1 [-1]
  ask n-of grupos patches [
    sprout 1 [
      set size 1
      set tipo 1
      ;set neighbors-list n-values 1 [-1]
      set color first available-colors
      set available-colors butfirst available-colors
      ;set level-of-adoption 0
    ]
  ]
  ask patches [;with [tipo != 2] [
    set pcolor [color] of min-one-of turtles [distance myself] ;Paints the farm with the
    color of the house
  ]
]

```

```
set tutor-related [who] of min-one-of turtles [distance myself] ;Makes a tag to the patch
```

```
set own_color pcolor
```

```
set Prob_fraude random-float 1;; Prob Fraude: Probabilidad de incurrir fraude
```

```
ifelse Perc_est_severidad_sancion = "Baja"
```

```
[ set Perc_severidad random-float 0.5] ;; Perc_severidad: Percepcion de severidad]
```

```
[ set Perc_severidad 0.5 + random-float 0.5]
```

```
ifelse Percepcion_Est_dificult_act = "Baja"
```

```
[ set Perc_dificul_exam 2 + random-float 3;; Percepcion de la dificultad de las tareas, es la m para la triangular
```

```
set Nota_est 9.9] ;; Perc_severidad: Percepcion de severidad
```

```
[ set Perc_dificul_exam random-float 5;; Percepcion de la dificultad de las tareas, es la m para la triangular
```

```
set Nota_est 9.9] ;; Perc_severidad: Percepcion de severidad]
```

```
]
```

```
ask turtles [
```

```
set Num_estudiantes count patches with [pcolor = [pcolor] of myself]
```

```
;set Num_estudiantes count patches with [pcolor = tutor-related ]
```

```
set color black
```

```
ifelse Percepcion_fraude_Profesor = "Laxo"
```

```
[ set Perc_gravedad_fraude random-float 0.5]
```

```
[ set Perc_gravedad_fraude 0.5 + random-float 0.5]
```

```
set Frecuencia_tareas Actividades_Por_asignatura
```

```
]
```

```
end
```

```
to-report triang [a b c]
```

```
let temp random-float 1
```

```
ifelse temp < (c - a) / (b - a)
```

```
[ report a + sqrt (temp * (b - a) * (c - a)) ]
```

```
[ report b - sqrt ((1 - temp) * (b - a) * (b - c))]
```

```
end
```