CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS

DEL INSTITUTO POLITÉCNICO NACIONAL

Unidad Zacatenco

Departamento de Matemática Educativa

El estudio de los conceptos de distribución de probabilidad y regresión lineal con el uso de tecnologías digitales con estudiantes de bachillerato

Tesis que presenta

Miguel Cerón Villegas

Para obtener el grado de

Maestro en Ciencias

En la especialidad de

Matemática Educativa

Director de la tesis: **Dr. Luz Manuel Santos Trigo**

Ciudad de México Octubre, 2017

Agradecimiento especial al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo brindado al otorgarme la beca para la realización de mis estudios de Maestría en Ciencias en el Departamento de Matemática Educativa del CINVESTAV-IPN

Becario: 599142

Agradecimientos

Al Dr. Luz Manuel Santos Trigo, por su apoyo, orientación y dedicación, durante el transcurso del desarrollo de esta investigación.

A mis sinodales, Dr. François Pluvinage y Dr. Roberto Ávila Antuna por sus recomendaciones y aportaciones para la mejora del trabajo.

A los doctores, que me impartieron los cursos de maestría, por compartir sus recomendaciones, experiencias y sobre todo conocimientos.

A mis compañeros de maestría, Alexandra, Alfredo, Ernesto, Javier, Mariana, Marcos, Maribel, Mayra, Salaj y Yesenia, por ser cómplices y participes de esta aventura.

A mis colegas de profesión, en especial a Herman, Fabián, Ricardo y Arturo, por ser amigos incondicionales.

A mis compañeros y amigos de Instituto Politécnico Nacional por ser siempre una referencia.

A todo el equipo de resolución de problemas, Adrián, William, Dany, Isaid, Julio y Yesenia Basaldúa, por ser excelentes compañeros, amigos y por ayudarme a ver las matemáticas desde otra perspectiva.

Dedicatoria

El presente trabajo es el resultado del esfuerzo de toda mi familia, mi padre Agustín es mi ejemplo a seguir, mi madre Lupita por ser mi más grande amor, mis hermanos Hugo y Elizabeth por ser los mejores críticos y compañeros de vida, a mi compadre Óscar por ser quién me inspira y a Matías el más pequeño del clan porque representa la diversión y el futuro de nosotros. Para todos mi eterna gratitud.

Tabla de contenido

Resumen 1
Abstract 2

Capítulo 1. Antecedentes y problema de investigación 3
1.1 La enseñanza de la probabilidad y la estadística 3
1.2 La enseñanza de la probabilidad y estadística con el uso de tecnología 6
1.3 El problema de investigación 9
1.4 Preguntas de investigación 10

Capítulo 2. Revisión de literatura 13
2.1 Introducción 13
2.2 La solución de tareas de probabilidad y estadística en un ambiente de resolución de problemas 13
2.2.1. La resolución de problemas según Polya 13
2.2.2. La resolución de problemas según Schoenfeld 14
2.2.3. La resolución de problemas y el uso de la tecnología digital 16
2.3 Las estrategias en la resolución de problemas 18
2.4 La importancia del estudio de las distribuciones de probabilidad 19
2.5 La importancia de entender la inferencia estadística 21
2.6 Herramientas tecnológicas para la enseñanza de estadística y probabilidad 23
2.6.1 Software educativo 23
2.6.2 Hojas de cálculo 24
2.6.3 Aplicaciones autónomas /Archivo en GeoGebra 24
2.6.4 Internet 25

Capítulo 3. Elementos metodológicos, diseño del estudio y procedimientos 27
3.1 Introducción 27
3.2 Características de una investigación cualitativa 27
3.3 Los participantes 27
3.4 Sesiones de trabajo y actividades 28
3.4.1 Sesiones de trabajo 28
3.4.2 Diseño de las actividades 29
3.4.3 Primera actividad – Cuestionario sobre evento aleatorio 31
3.4.4 Segunda actividad – Lanzamiento de dados 31
3.4.5 Tercera actividad – la población joven por mediante un ajuste de línea recta 32
3.5 Instrumentos de recolección de información 35
3.6 Estudio estadístico de las actividades 36
3.6.1 Estudio del lanzamiento de los dados 36
3.6.2 Estudio de la población joven por ajuste de una línea recta 39
3.7 Análisis de datos 42

Capítulo 4. Análisis de Datos 47
4.1 Primera actividad – Análisis del cuestionario sobre el evento aleatorio 47
4.2 Segunda actividad – Lanzamiento de dados 52
4.2.1 Análisis de la segunda actividad 72
4.3 Tercera actividad – Ajuste de una línea recta 75
4.3.1 Análisis de la tercera actividad 84

Capítulo 5. Conclusiones 87
5.1 Conclusiones 87
5.2 Reflexiones finales 91
Revisión bibliográfica 93
Anexo 1. Actividades de la investigación 99
Anexo 2. Tablas de registros de los dados inventados por los alumnos 111

Resumen

El propósito de este estudio fue analizar y documentar cómo el uso de tecnologías digitales ofrece oportunidades para que los estudiantes comprendan y exploren el comportamiento de eventos aleatorios. Para comprender los conceptos de la función de distribución de probabilidad y la regresión lineal, los participantes realizaron actividades que involucraron la recolección de datos de eventos reales, proponiendo resultados individuales y usando una aplicación para simular un evento. Las preguntas de investigación que guiaron el desarrollo del estudio fueron:

(1) ¿De qué manera la simulación de un experimento aleatorio con el uso de GeoGebra y Excel influye en el desarrollo de la comprensión del concepto de función de distribución de probabilidad en estudiantes a nivel de bachillerato?

(2) ¿Cómo influye el uso de internet y GeoGebra en la comprensión que muestran estudiantes de bachillerato en el estudio del concepto de regresión lineal?

13 estudiantes de secundaria participaron en el estudio. Ellos trabajaron en actividades que involucraron registrar la suma de números que aparecen al lanzar dos dados y analizar el comportamiento de los datos (porcentaje de la población joven) a través de un modelo lineal. Los resultados indicaron que fueron capaces de interpretar y entender el comportamiento de la distribución de frecuencias simulando el evento mediante el aumento del número de ensayos y encontraron coeficientes asociados con la pendiente y la ordenada para ajustar una línea que mejor se ajuste a los datos.

Abstract

The purpose of this study was to analyze and document how the use of digital technologies provides opportunities for students to understand and explore the behavior of random events. To comprehend the concepts of the probability distribution function and linear regression the participants relied on activities that involved gathering data of real events, proposing individual results and using an applet to simulate the events. The research questions that guided the development of the study were:

(1) To what extent the use of GeoGebra to simulate the random event results influence the participants comprehension of the concept of distribution of probability?

(2) How the use of internet and GeoGebra helped the participants to make sense of data and propose and justify the use of linear regression model.

13 high school students participated in the study. They worked on activities that involved recording the sum of numbers that appears while throwing two dice and analyzing the behavior of data (percentage of young population) through a lineal model. Results indicated that they were able to interpret and understand the behavior of the distribution of frequencies by simulating the event by increasing the numbers of trials and they found coefficients associated with slope and the ordinate to adjust a line the best fit to data.

Capítulo 1. Antecedentes y problema de investigación

1.1 La enseñanza de la probabilidad y la estadística

El estudio de la probabilidad y la estadística en todos los niveles educativos está adquiriendo más atención, ya que proporciona a los estudiantes herramientas, ideas y disposiciones para representar, explorar e interpretar la información sobre eventos que se produce y difunde cotidianamente.

El proporcionar argumentos basados en evidencia y evaluar críticamente las afirmaciones o conclusiones a partir de analizar datos, son habilidades importantes que todos los ciudadanos deberían construir. Como una respuesta a la necesidad de mejorar la capacidad de los estudiantes para pensar estadísticamente, la alfabetización estadística y el razonamiento forman parte de los currículos escolares y universitarios de muchos países. Como consecuencia, la educación estadística emerge como un campo de investigación y de desarrollo curricular (Ben-Zvi & Bakar, 2016).

La omnipresencia de los datos en la vida cotidiana, por ejemplo, el producto interno bruto de un país, la población infantil de un estado, las preferencias políticas de una comunidad, las calorías que un adulto consume, etcétera, se pueden consultar fácilmente con dispositivos y desarrollo digital. En la actualidad los ciudadanos cotidianamente encuentran información cuantitativa en los medios de comunicación. En el lugar de trabajo, los datos son vitales para el control de calidad, el monitoreo y la mejora de la productividad, y la anticipación de los problemas (Bakker, Kent, Noss, & Hoyles, 2009). Con los datos cada vez más usados para agregar o implicar credibilidad, las escuelas deben preparar a sus estudiantes para crear y evaluar críticamente las afirmaciones basadas en datos. El progreso en la comprensión de la enseñanza y el aprendizaje del razonamiento estadístico y la disponibilidad de herramientas digitales para el aprendizaje de las estadísticas han permitido a un campo relativamente reciente de la educación estadística integrarse y aprovechar fácilmente estos avances (Garfield & Ben-Zvi, 2008).

Ben-Zvi (2014) señala que el objetivo de enseñar estadística es promover que los estudiantes desarrollen conocimientos del dominio o área y tengan la capacidad de razonar estadísticamente. La alfabetización estadística es la capacidad de interpretar, evaluar críticamente y comunicar información y mensajes estadísticos. El razonamiento estadístico es la forma en que las personas representan y exploran las ideas estadísticas y dan sentido a la información estadística que induzca información cuantitativa. Este tipo de pensamiento es la forma en que la información se ve, se procesa y se convierte en pasos de acción, es una filosofía de pensamiento, no una forma de realizar cálculos matemáticos. El pensamiento estadístico utiliza el concepto de que toda actividad consiste en un conjunto de pasos interconectados que deben complementarse y completarse para lograr una meta planteada, además puede implicar la conexión de un concepto a otro (por ejemplo, el centro y la dispersión) o puede combinar ideas acerca de los datos y el azar. Además, este razonamiento también puede ayudar a entender y explicar los procesos estadísticos y ser capaz de interpretar los resultados obtenidos.

El azar se encuentra presente en nuestras vidas, en entornos donde aparecen nociones de incertidumbre, riesgo y probabilidad, por ejemplo, el pronóstico del tiempo, un diagnóstico médico, estudio de la posibilidad de tomar un seguro de vida o efectuar una inversión, etc. Las personas deben analizar situaciones con estos elementos, tomar decisiones, emitir juicios sobre relación entre sucesos o efectuar inferencias y predicciones (Gigerenzer, 2002). En estas situaciones la probabilidad no es una propiedad física tangible u objetiva de los sucesos que nos afectan, como el peso, color, superficie o temperatura, sino una manera de cuantificar la ocurrencia de un evento.

Begg (1997) menciona que la probabilidad y la estadística son un buen vehículo para alcanzar las capacidades de comunicación, tratamiento de la información, resolución de problemas, uso de computadoras y trabajo cooperativo, a las que se da gran importancia en los nuevos currículos. Además, la probabilidad y la estadística no requieren técnicas matemáticas complicadas. Sus aplicaciones, proporcionan una buena oportunidad para mostrar a los estudiantes la utilidad de la matemática para resolver problemas reales, siempre que su enseñanza se lleve a cabo mediante la participación activa de los estudiantes, enfatizando la experimentación y la resolución de problemas.

Shaughnessy (1992), refiere que existe un vínculo entre la resolución de problemas y la probabilidad y estadística. En la enseñanza y el aprendizaje de la probabilidad y estadística es necesario la construcción de modelos que representen fenómenos físicos, el desarrollo y el uso de estrategias (tales como la simulación y las técnicas de conteo), y la comparación y evaluación de varios y diferentes acercamientos a los problemas. En este sentido, enseñar probabilidad y estadística es enseñar la resolución de problemas, aunque en el dominio de un contenido particular. También menciona que los antecedentes de los maestros tanto en la probabilidad y estadística como en la resolución de problemas, son débiles o no existen, debido a que los programas de preparación de maestros no han incluido sistemáticamente ni la probabilidad y estadística ni la resolución de problemas en la formación de los profesores.

El uso de la tecnología ha dado lugar a numerosos cambios en la enseñanza de la estadística, si bien es importante que los alumnos puedan realizar algunas actividades de simulación con apoyo de material concreto, como monedas, dados, ruletas, con tablas de números aleatorios y calculadoras, es realmente la computadora la que proporciona una mayor potencia de simulación, construcción de modelos y exploración de los mismos, ya sea mediante el uso de software estadístico[1], software educativo[2], con la manipulación de aplicaciones autónomas[3] o con las hojas de cálculo electrónicas, con las cuales se puede procesar una gran cantidad de información, para después poder ser analizada y proceder a realizar su representación gráfica (Garfield &Ben-Zvi, 2008).

En el nivel bachillerato, cada vez hay más interés para ayudar a los estudiantes a desarrollar una comprensión y familiaridad con el análisis exploratorio de datos (EDA, por sus siglas en inglés), el modelado y la simulación en lugar de enseñarles un conjunto de habilidades y procedimientos separados. Se da énfasis en proyectos, combinado con esfuerzos cada vez mayores para integrar la tecnología en la reforma del aprendizaje (Biehler, Ben-Zvi, Bakker, & Makar, 2013).

---

[1] Programa informático diseñado especialmente para resolver problemas en el área de la estadística.
[2] Herramienta pedagógica o de enseñanza que por sus características está vinculado a la educación.
[3] Componente de software que realiza una función específica y pueden ayudar a los estudiantes a explorar conceptos en un entorno visual, interactivo y dinámico.

Batanero (2014) señala que con la creciente importancia de las estadísticas en la escuela y el progreso de la tecnología con fácil acceso a la simulación, hoy en día existe un interés creciente en un acercamiento experimental al estudio de la probabilidad como un límite de frecuencias que mantienen un comportamiento definido. También se observa un cambio en la forma en que se enseña la probabilidad desde un enfoque basado en fórmulas hasta una introducción experiencial moderna donde el énfasis está en la experimentación de sucesos probabilísticos. Además, menciona que se debe alentar a los estudiantes a realizar experimentos o simulaciones aleatorios, formular predicciones sobre la tendencia de resultados en una serie de estos experimentos, recopilar y analizar datos para probar sus conjeturas y justificar sus conclusiones sobre la base de estos datos. Batanero (2014) indica que siguiendo este enfoque se puede mostrar a los estudiantes que la probabilidad es inseparable de las estadísticas, y viceversa, tal como se reconoce en los planes de estudios.

1.2 La enseñanza de la probabilidad y estadística con el uso de tecnología

Garfield y Ben-Zvi (2008), enuncian que: 1) la enseñanza de la probabilidad y estadística en la actualidad es difícil imaginar sin el uso de alguna forma de tecnología. 2) En las clases de estas asignaturas en el presente, algunos profesores suelen hacer uso de una computadora con una pantalla para proyectar su clase, además de que con mayor frecuencia en los centros educativos existen laboratorios de computo en donde los estudiantes trabajan en sus propios equipos. 3) Algunos alumnos en sus escuelas pueden usar una computadora portátil o de escritorio. Una propuesta de la enseñanza de la probabilidad y la estadística en nuestros días se puede organizar con el uso del internet, en forma de un curso basado en la Web con conferencias grabadas en video, discusiones interactivas, proyectos de colaboración, texto electrónico y materiales de evaluación. Estos investigadores, señalan que estas herramientas empleadas de manera coordinada, conducen a una mejora en el aprendizaje de la probabilidad y la estadística.

Santos-Trigo (2017) argumenta que, la entrada al mundo digital ha sido repentina y no tiene vuelta atrás; los teléfonos móviles inteligentes, las tabletas y laptops son herramientas imprescindibles en muchas tareas que realiza el individuo. La irrupción tecnológica ha traído claros beneficios a todos aquellos que disponen de un dispositivo digital; pero también ha generado retos importantes en la enseñanza; en consecuencia, se necesita más información sobre el diseño de tareas matemáticas cuando se utilizan nuevas tecnologías, debido a que el uso de nuevas herramientas requiere de una nueva didáctica. Pero, ¿cómo la tecnología puede apoyar el proceso de aprendizaje de los estudiantes?

Los estudiantes manifiestan ser capaces de aprender a usar diferentes tipos de software en un mismo curso, pero los maestros también deben poseer una comprensión de los alcances de esas tecnologías, al menos para proporcionar las ayudas que genera su uso, y para proporcionar orientación sobre cuándo utilizarlas, ya que si se usan adecuadamente, se puede simplificar los cálculos aritméticos, además de realizar representaciones gráficas en un menor tiempo, con una alta precisión y con muy pocos errores.

La disponibilidad de herramientas tecnológicas cada vez más poderosas y versátiles han posibilitado que tareas complejas sean realizadas más fácilmente (Leung, 2013). Al respecto, el Consejo Nacional de Profesores de Matemáticas (NCTM por sus siglas en inglés, 2008) resalta la importancia del uso de la tecnología digital como potenciador del aprendizaje matemático:

> La tecnología es una herramienta esencial para el aprendizaje de las matemáticas en el siglo XXI y tanto los profesores como los estudiantes deben tener acceso regular a las tecnologías que apoyan las actividades de dar sentido matemático, razonamiento, resolución de problemas y comunicación (p.1).

Garfield y Ben-Zvi (2007) exponen que la tecnología también ha ampliado las formas realizar gráficas para la visualización de datos, esto puede para ayudar a los estudiantes a explorar y analizar datos. Permitiéndoles enfocarse en la interpretación de los resultados y comprensión de conceptos en lugar de centrarse solo en los aspectos procedimentales. La interactividad, con el uso de un deslizador (herramienta en un software que permite elegir valores en un rango definido) puede ayudar a los estudiantes a centrarse en los efectos de cambiar valores en el cálculo sin la necesidad de recalcular nuevamente los términos y así potenciar el pensamiento matemático de los estudiantes.

La tecnología puede desempeñar un papel importante en la mejora de la capacidad de los estudiantes para estudiar los fenómenos aleatorios y los conceptos estadísticos, a través de la visualización y el diseño de simulaciones (Chance & Rossman, 2006). ¿Cómo pueden las simulaciones ser empleadas como herramienta pedagógica? Por ejemplo, estas herramientas permiten a los estudiantes responder lo que sucede si esto se repite un gran número de veces a través de la observación directa.

Shaughnessy (2007) sugiere que los estudiantes están mejor capacitados para desarrollar intuiciones correctas cuando son introducidos en el estudio de la probabilidad usando simulaciones de fenómenos aleatorios. La simulación permite condensar el experimento en el tiempo y en el espacio y operar con el experimento simulado para obtener observaciones y conclusiones válidas sobre fenómeno en estudio. Además, como lo señala Mills (2004) en ocasiones, las simulaciones pueden jugar un rol significativo en el desarrollo de las habilidades de los estudiantes cuando se estudian procesos aleatorios y conceptos estadísticos, ya que ayuda al alumno a conocer las diferencias entre la probabilidad experimental y el acercamiento teórico.

El NCTM resalta explícitamente la importancia de la simulación en la probabilidad y la estadística:

> Los estudiantes de nivel bachillerato deberían poder avanzar desde situaciones en las que la probabilidad de un suceso se puede determinar fácilmente, a situaciones en las que la toma de muestras y las simulaciones les ayudan a cuantificar la probabilidad de un resultado incierto. (NCTM, 2011, pág 328)

Konold y Kazak (2008) sugieren estudiar a los eventos y el muestreo como procesos aleatorios a partir del desarrollo de ideas intuitivas de los estudiantes sobre probabilidad, poblaciones, muestras, variabilidad muestral, distribuciones muestrales y efecto del tamaño de muestra y así desarrollar una concepción aleatoria de la probabilidad y la inferencia estadística desde edades tempranas. Este acercamiento a los conceptos de inferencia desde una perspectiva informal requiere del uso de herramientas tecnológicas con amplio potencial de representaciones visuales dinámicas para generar imágenes de estos conceptos en los estudiantes. De esta forma cuando se emiten juicios o se toman decisiones en entornos caracterizados por la incertidumbre, el pensamiento probabilístico se convierte en una herramienta para que los estudiantes comprendan y puedan predecir hechos o nuevos comportamientos.

Ben-Zvi y Bakar (2016), sugieren que los entornos de aprendizaje incorporen datos auténticos (en muchos casos, datos reales, no sólo que se ajusten a la realidad) y que los estudiantes recolecten, organicen y representen, analicen e infieran, comuniquen y discutan hallazgos e ideas, ya que la tecnología hoy día permite aplicar la estadística con gran facilidad, cobra mayor importancia las actividades interpretativas que el cálculo numérico.

La tecnología facilita la discusión de problemas más complejos, analizando un conjunto de datos que puede ser numeroso, el cual a menudo se accede desde internet, buscando que la fuente que lo proporcione sea una institución confiable. Ahora tenemos la facilidad de hacer que los estudiantes analicen datos reales, dándoles una mejor idea de lo que hacen los estadísticos, haciendo que ellos pasen por el proceso de recopilar, analizar y sacar sus propias conclusiones, contestando una serie de preguntas, de esta forma se puede hacer que comprendan y experimenten mejor la práctica de la estadística (Ben-Zvi, 2004a). Surgen preguntas como las que realiza Santos-Trigo (2017), ¿cómo construir escenarios de aprendizaje que extiendan las actividades escolares que se llevan a cabo en la enseñanza formal? ¿qué materiales se deben diseñar y qué tipo de problemas y actividades deben abordar los estudiantes con el uso de diversas tecnologías?

1.3 El problema de investigación

El problema de investigación se centra en analizar y documentar cómo el uso sistemático y coordinado de las tecnologías digitales en un ambiente de resolución de problemas, permiten a los estudiantes obtener datos reales, graficar, explorar, interpretar e inferir resultados. Se intenta registrar las ideas que muestran los alumnos en la comprensión de un evento aleatorio y las habilidades estadísticas (recolectar, graficar, analizar e interpretar la información obtenida). Además, se busca que los participantes registren sus conclusiones y le den sentido a las mismas a partir de la obtención de los resultados estadísticos.

Así, en un ambiente de resolución de problemas dentro del aula, los estudiantes pueden utilizar las tecnologías digitales como GeoGebra, hojas de cálculo, repositorio de datos y materiales (Classroom) e internet que además les permitan revisar conceptos y extender la discusión de contenidos matemáticos fuera del aula. Santos-Trigo y Reyes-Martínez (2014), mencionan que el uso coordinado de estas tecnologías puede ofrecer distintas oportunidades a los estudiantes para desarrollar el pensamiento matemático; el cual tiene como rasgos distintivos el visualizar, representar, buscar patrones, conjeturar y probar, como algunas de las actividades que los estudiantes deben desarrollar en sus experiencias de aprendizaje.

"En una época de tecnología y desarrollos digitales resulta importante preguntar: ¿Qué oportunidades ofrece al individuo el uso de tales herramientas en la construcción de conocimiento y formas de pensar compatibles con una cultura científica y estrategias para formular y resolver problemas?, en general, la mayoría de las instituciones educativas mantienen los mismos cursos y contenidos que se han ofrecido desde principios del siglo pasado y reproducen las mismas prácticas de enseñanza. En ese ambiente de aprendizaje, un profesor, normalmente ubicado al frente del salón, explica los conceptos o contenidos a los estudiantes, quienes escuchan y pocas veces preguntan o intercambian ideas entre ellos mismos durante la clase. Es común también que exista un libro de texto para seguir las explicaciones del profesor y que los estudiantes resuelvan exámenes diseñados por el profesor para evaluar su desempeño", Santos-Trigo (2016).

Como consecuencia, un factor que reclama la investigación sobre el impacto que aportan las herramientas digitales en el aprendizaje de la probabilidad y la estadística de los estudiantes es la premisa de que diferentes herramientas digitales favorecen el desarrollo del pensamiento estadístico; así como uso de secuencias didácticas que empleen el método inquisitivo, el cual pueda favorecer a que los estudiantes tengan acercamientos a conceptos estadísticos, dentro de un ambiente de resolución de problemas.

1.4 Preguntas de investigación

A partir de la revisión de la literatura se diseñó un conjunto de actividades para que un grupo de estudiantes de bachillerato haciendo uso de diversas herramientas digitales, analizaran el comportamiento de un evento aleatorio, para llegar a la comprensión del concepto de distribución de probabilidad, además utilizando un conjunto de datos oficiales tomados del portal de la OCDE[4] (Organización para la cooperación y el desarrollo económico) y el uso de GeoGebra, identificar la comprensión del concepto de regresión lineal. Así, se formulan las siguientes preguntas que guiaron el desarrollo del estudio:

¿De qué manera la simulación de un experimento aleatorio con el uso de GeoGebra y Excel influye en el desarrollo de la comprensión del concepto de función de distribución de probabilidad en estudiantes a nivel de bachillerato?

El propósito de esta pregunta es identificar la forma en cómo la simulación de un experimento aleatorio construida en GeoGebra y las gráficas en Excel influyen en el desarrollo de la comprensión de un evento aleatorio, contrastando los resultados obtenidos al realizar el experimento, al inventar resultados y al simular el evento respondiendo una serie de preguntas que orientaron a los alumnos a conectar los conceptos de la distribución de frecuencias con la distribución de probabilidades.

¿Cómo influye el uso de internet y GeoGebra en la comprensión que muestran estudiantes de bachillerato en el estudio del concepto de regresión lineal?

Con esta pregunta, se busca documentar el tipo de procedimientos que emplean los alumnos al hacer uso de datos oficiales del sitio de internet de la OCDE y el uso de la tecnología digital al estudiar el concepto de regresión lineal para realizar inferencias estadísticas y con ello dar respuesta a las preguntas planteadas.

[4] https://data.oecd.org

Capítulo 2. Revisión de literatura

2.1 Introducción

En este capítulo se presentan y discuten algunos resultados de investigaciones que se han publicado recientemente en torno al estudio y aprendizaje de la probabilidad y estadística. Inicialmente, se aborda un marco conceptual basado en la resolución de problemas que ayudó a estructurar el desarrollo del estudio. Se describe la importancia del estudio de las distribuciones de probabilidad, así como la relevancia que tiene el análisis de datos bivariados con ayuda de una recta de regresión lineal, para la inferencia de resultados. Se expone la importancia de las tecnologías digitales en la enseñanza de la probabilidad y estadística; y finalmente se resalta el uso de datos reales en la solución de problemas estadísticos.

2.2 La solución de tareas de probabilidad y estadística en un ambiente de resolución de problemas

En el contexto de la educación matemática han surgido propuestas curriculares y modelo de enseñanza desde las cuales se concibe, piensa y desarrolla la resolución de problemas. A continuación, se identifican los elementos importantes de la resolución de problemas que sustentan el desarrollo de este estudio.

2.2.1. La resolución de problemas según Polya

Un problema existe cuando se comprende el enunciado y la meta u objetivo, pero se desconoce el camino como resolverlo, Polya (1962). El profesor debe ofrecer una ayuda discreta y tratar de comprender lo que pasa por la mente del estudiante, pues solo así podrá brindarle el apoyo necesario para que pueda conseguir autonomía y lograr sus objetivos. Algunas preguntas que Polya propone que se le planteen al estudiante para que se concentre en la incógnita de algún problema son: ¿Qué se requiere?; ¿qué quiere usted determinar?; ¿qué se le pide a usted que encuentre? El fin es esclarecer el problema a partir de preguntas sencillas.

Polya (1945) identifica cuatro fases necesarias para resolver un problema:

- En la fase de comprensión del problema un componente fundamental es que el estudiante tenga el deseo de resolver el problema, no basta únicamente con comprenderlo. Es necesario comprender el problema: ¿cuál es la incógnita?, ¿cuáles son los datos y las condiciones?
- La fase del diseño de un plan es primordial para la solución de un problema, porque aquí se construye la idea que dará sustento al cómo proceder, y para que esto suceda el estudiante debe basarse en sus experiencias y conocimientos adquiridos previamente, ¿conoce un problema relacionado con éste?, ¿conoce algún teorema que le pueda ser útil?, ¿podría enunciar el problema de otra forma?, ¿Ha empleado todos los datos?
- La fase de implementación del plan, implica conocimientos ya adquiridos, buenos hábitos de pensamiento, concentración y paciencia; para comprobar cada uno de los pasos, ¿puede usted ver que el paso es correcto?
- La visión retrospectiva es, la fase en donde el estudiante verifica el resultado, examinando el resultado y el procedimiento que lo condujo a él, consolidando sus conocimientos y desarrollando sus aptitudes para resolver problemas, realizando extensiones y conexiones con otros temas.

2.2.2. La resolución de problemas según Schoenfeld

Schoenfeld (1985) en su libro *Mathematical Problem Solving*, considera que las estrategias que plantea Polya para la resolución no son suficientes y propone un marco conceptual basado en la resolución de problemas, que permita entender cómo intentan los estudiantes resolver problemas y a partir de ello, qué actividades se pueden formular para ayudarlos. El marco consta de cuatro dimensiones que analizan el proceso que sigue el individuo al resolver problemas:

- Dominio del conocimiento o recursos.
- Estrategias cognitivas o métodos heurísticos.
- Estrategias metacognitivas.

- Sistemas de creencias.

Dominio del conocimiento o recursos básicos: son los conocimientos previos que posee la persona, se refiere, entre otros, a conceptos, fórmulas, algoritmos, y en general todas las nociones que se considere necesario saber para enfrentar un problema. Un elemento clave a tener presente es el de ver si el estudiante tiene ciertos estereotipos o recursos defectuosos o mal aprendidos. Schoenfeld (1985) menciona que una parte de los recursos es el conocimiento informal e intuitivo, el cuál es la forma de interpretar conocimiento nuevo por parte de un individuo. Esta interpretación (no siempre correcta) está fundamentada en conocimientos asociados a su vida diaria o en conocimientos previos. Este tipo de conocimiento se debe tomar en cuenta en los escenarios de enseñanza, pues en algunas ocasiones es favorable usar las ideas intuitivas del individuo para desarrollar un conocimiento formal, en otras ocasiones es necesario confrontar y desechar dichas ideas.

Estrategias cognitivas o métodos heurísticos: son las estrategias cognitivas que resultan importantes en el proceso de resolver problemas, también conocidos como métodos heurísticos; entre las cuales se incluyen las analogías, la introducción de elementos auxiliares en el problema, la argumentación por contradicción, la descomposición y recombinación de los datos, dibujar figuras, reformulación de problemas, emplear la reducción al absurdo, realizar pruebas indirectas, e incluso trabajar hacia atrás. En el estudio de la probabilidad y estadística existen las heurísticas de la representatividad, la disponibilidad y el sesgo de equiprobabilidad, que resultan de la realización de juicios intuitivos obtenidos bajo la incertidumbre y que más adelante se discuten.

Estrategias metacognitivas: también conocido como control, esta etapa ofrece la manera en que los individuos usan la información potencial a su disposición, es decir que controle su proceso, entendiendo de qué trata el problema, considerando varias formas de solución, seleccionando una específica, monitoreando de su proceso (incluyendo el tiempo), para verificar su utilidad y revisar que sea la estrategia adecuada.

Sistema de creencias: las creencias van a afectar la forma en la que un individuo se enfrenta a un problema matemático. Es la *visión matemática del mundo*, que posee la persona, es el conjunto de determinantes no necesariamente conscientes de la conducta de un individuo, acerca de uno mismo, sobre medio ambiente, sobre el tema y sobre las matemáticas.

En los estándares de la NCTM (2011), se proponen que cuatro actividades, si se desea resolver problemas de estadística: especificar el problema y planificar; recoger datos de una variedad de fuentes adecuadas; procesar y representar los datos; interpretar y discutir los resultados. Estas actividades se consideran que son cíclicas, ya que puede ser necesario refinar el enfoque inicial para resolver un problema y repetir el proceso de nuevo. En este sentido las estrategias propuestas por Polya y Schoenfeld, se encuentran en sincronía con las actividades propuestas en la presente investigación, pero además es necesario considerar el uso de las tecnologías digitales, como se muestra a continuación.

2.2.3. La resolución de problemas y el uso de la tecnología digital

Dado que las actividades propuestas en el presente trabajo involucraron el uso de herramientas digitales, se tomó el marco propuesto por los investigadores Santos-Trigo y Camacho-Machin (2009). Estos investigadores proponen una forma de caracterizar episodios, en términos de organización y estructura, del proceso de resolución de problemas partiendo de las etapas identificadas por Polya (1945).

La resolución de problemas constituye una parte integral de todo el aprendizaje de las matemáticas, además de ser un marco que se ajusta a las necesidades de este trabajo, debido a que como comenta Santos-Trigo (2014), permite caracterizar el desarrollo del pensamiento matemático en los estudiantes al poder, analizar los recursos básicos con los que cuentan, las estrategias cognitivas o heurísticas, y por último las creencias. Siguiendo esta perspectiva, se sugiere que la resolución de problemas debe:

- Ayudar a los estudiantes a desarrollar estrategias de resolución de problemas con cierto grado de especificidad y que relacionen de manera precisa clases particulares de problemas donde se aplican.
- Desarrollar formas de robustecer las creencias de los estudiantes sobre la naturaleza de las matemáticas, la resolución de problemas, y sobre sus propias competencias o formas de interactuar con situaciones matemáticas.

Una característica esencial de este marco es que incorpora el uso de herramientas digitales como un elemento central en el proceso de resolución de problemas, en donde se proponen los siguientes episodios:

- Comprensión del problema

En este episodio se requiere fomentar un método inquisitivo para que los estudiantes piensen el problema en términos de cuestiones relevantes que conduzcan a la exploración y representación de relaciones matemáticas. Algunos de los objetivos de esta etapa son:

> *Identificar los objetos matemáticos* involucrados en el problema y proponer formas de representarlos. Aquí, el uso de herramientas computacionales juega un papel importante.
>
> *Evaluar la pertinencia del problema*, es decir, verificar que el problema tenga sentido y sea factible.
>
> *Relacionar las condiciones del problema con los objetos matemáticos*. En otras palabras, se requiere explicitar la dependencia entre dichos objetos involucrados en el enunciado del problema.

- Implementación de un plan de solución

En este episodio es importante que se implementen diferentes acercamientos a la solución del problema; se requiere que las condiciones del problema, y su relación con los objetos matemáticos, se modelen desde diferentes perspectivas. Los acercamientos a la solución pueden ser visuales-dinámicos, numéricos-empíricos, geométricos, algebraicos, etc. Al realizar estos acercamientos se puede hacer uso de herramientas digitales que favorezcan el acceso a los recursos necesarios para resolver el problema.

- Búsqueda de patrones y una solución general

El quehacer matemático puede ser caracterizado como la actividad de encontrar y examinar patrones entre los diferentes acercamientos a la solución de un problema matemático. En consecuencia, en esta etapa, el uso de la tecnología juega un papel determinante; los patrones numéricos pueden ser explicitados mediante el uso de una hoja de cálculo, un análisis algebraico puede ser implementado vía calculadora simbólica, los patrones geométricos pueden ser visibles a través de invariantes en configuraciones a través de una aplicación o tecnología digital, o como en el presente trabajo a través de la simulación de un evento aleatorio. El reconocimiento de patrones, en los distintos acercamientos, puede conducir a la solución general del problema.

- Conexiones y extensiones

En este episodio se favorece el tratamiento de los acercamientos a la solución alrededor de las propiedades. Es decir, se requiere que los conceptos matemáticos utilizados en los diversos acercamientos (empírico, dinámico, geométrico, algebraico, etc.) se expliciten y relacionen. Y la tecnología también puede ayudar a crear vínculos, buscando generalidades y métodos utilizados en los resultados y soluciones obtenidos.

2.3 Las estrategias en la resolución de problemas

En la resolución de un problema se pueden seguir procedimientos algorítmicos. Los algoritmos (viene de guarismo = cantidad y es una alteración por influjo del griego "arithmós") son procedimientos que aplican en la solución; por ejemplo, un algoritmo son las reglas para realizar una división cualquiera de dos números, es decir lo que haríamos para dividir 240/30. Estas reglas nos darán un resultado basado en la aplicación de un procedimiento. Las estrategias heurísticas en cambio, (de "heuriskó", yo descubro, en griego) (Corominas, 1973) son estrategias que proveen ayuda en la solución de un problema, pero no garantiza el éxito o resolución del problema.

Kahneman y Tversky (1973) y Lecoutre (1992), identifican tres heurísticas que obedecen a juicios intuitivos que se obtienen bajo la incertidumbre. Estos son la representatividad, la disponibilidad y el sesgo de equiprobabilidad (es decir, que dos o más sucesos tienen la misma probabilidad de ocurrir). Aunque la intuición heurística se distingue de los procesos de razonamiento formativo por pautas de juicios sesgados, los heurísticos en sí mismos son procedimientos de estimación que de ningún modo son irracionales. Son respuestas intuitivas normales, no solo para los problemas de alta complejidad, sino para las más simples cuestiones de verosimilitud, frecuencia y predicción. Además, mencionan que los procesos de juicios intuitivos no solo son más simples de los que exigían los modelos racionales, sino que eran categóricamente de una clase diferente.

Para estos investigadores, con la heurística de la representatividad las personas estiman las probabilidades de los eventos con base a lo bien que un resultado representa cierto aspecto de su población original, esta heurística también se utiliza para explicar el *efecto contrario a los valores recientes*, también conocida como la “falacia del jugador”, en donde los sujetos tienden a creer que después de una serie de soles, al lanzar una moneda, es más probable que se obtengan águilas. La ley de los grandes números garantiza que las muestras grandes sean representativas de la población de la que se han extraído; pero no las muestras pequeñas. Sin embargo, la intuición de las personas también les hace pensar que las pequeñas cantidades se regirán por las leyes de los grandes números, pero esto no es cierto.

Mientras que la heurística de disponibilidad se refiere a la tendencia de individuo a determinar qué probabilidad hay de que un suceso se dé o no. Cuando más reciente es un suceso, parecerá más frecuente y probable, cuanto más comprensible es la información, será más convincente y fácil de recordar, y cuanto más evidente resulta algo, parecerá ser más el motivo que lo genera. El sesgo o heurística de la disponibilidad es una tendencia a valorar las probabilidades con base en los ejemplos más sencillos que un individuo recuerda.

Lecoutre (1992), describe la creencia de los sujetos en la equiprobabilidad de todos los sucesos asociados a cualquier experimento aleatorio o donde no hay una simetría física. Para comprobar esta creencia, aplicó en sus experimentos un problema en el que se pregunta si al lanzar dos dados hay la misma probabilidad de obtener un 5 y un 6 que la de obtener dos veces un 5. A pesar de variar el contexto, el formato de la pregunta, la edad y la formación de los sujetos, los resultados siempre coinciden y demuestran la estabilidad de la creencia en que los dos resultados son equiprobables. Lecoutre defiende que ello no es debido a la falta de razonamiento combinatorio, sino a que los modelos combinatorios no se asocian fácilmente con las situaciones en que interviene "el azar". Los sujetos que muestran el sesgo de equiprobabilidad consideran que el resultado del experimento "depende del azar" y en consecuencia todos los posibles resultados son equiprobables.

2.4 La importancia del estudio de las distribuciones de probabilidad

Gran parte de la investigación sobre distribuciones de probabilidad surgió debido al consenso en la comunidad de educación estadística, de que su estudio sustenta una red de ideas estadísticas fundamentales, tales como variabilidad, muestreo e inferencia (Garfield, & Ben-Zvi, 2004; Pfannkuch, & Reading, 2006). Reading y Shaughnessy (2004), se enfocaron en investigar cómo los estudiantes razonan acerca de la distribución de probabilidad. Además, Chance, delMas y Garfield (2004) afirman que el conocimiento de la distribución de la probabilidad y la comprensión de los histogramas son requisitos previos necesarios para el aprendizaje y la comprensión de las distribuciones de frecuencias. Estos son algunos motivos por los que los resultados de la investigación en estadística proporcionan un lugar preponderante a la comprensión de las distribuciones, incluso en las formas más simples, debido a que su comprensión es mucho más compleja y difícil de lo que muchos profesores de estadística creen.

La representación gráfica de una distribución revela la variación de una variable cuantitativa, según Wild (2006), los estadísticos responden a la omnipresencia de la variabilidad en los datos. Los estudiantes analizan dos tipos principales de distribución en una clase introductoria de estadística. El primer tipo son las distribuciones de datos de muestra que los estudiantes aprenden a graficar, describir e interpretar. Éstas son distribuciones empíricas de un determinado fenómeno cuantificado. El segundo tipo de distribución es una distribución teórica, por ejemplo, distribuciones normales o binomiales, que son en realidad modelos de probabilidad.

Wild (2006) menciona que los dos tipos comparten muchas características comunes (por ejemplo, se pueden describir en términos de forma, centro y extensión), es importante ayudar a los estudiantes a distinguir entre ambas debido a la forma en que los usamos. La distinción que subyace a las distribuciones empírica versus teórica se relaciona con la variación. Al examinar una distribución empírica, el enfoque se centra en la descripción y la interpretación de los datos, y pensar en qué modelo puede explicar la variación de los datos. Las distribuciones teóricas son modelos que se ajustan a los datos, para ayudar a explicar, estimar o hacer predicciones sobre la variabilidad de los datos empíricos.

Los estudiantes a menudo ven y usan representaciones como ilustraciones y no como herramientas de razonamiento para aprender algo acerca de un conjunto de datos o para obtener información sobre un problema o contexto particular (Konold & Pollatsek, 2002). La investigación realizada por Bakker y Gravemeijer (2004), sobre la comprensión estadística de la distribución de los estudiantes, recomienda un cambio del enfoque de enseñanza que incluya dibujar diversos tipos de gráficas y desarrollar habilidades para dar sentido a los datos, detectar y descubrir patrones, para posteriormente confirmar o generar hipótesis, además de que sugieren que el análisis e interpretación constituye la base del razonamiento sobre las distribuciones.

Otros investigadores refieren que se deben desarrollar habilidades de decodificación visual, juicio y contexto como factores críticos para ayudar a los estudiantes a buscar el significado de las gráficas (Friel, Curcio, & Bright, 2001). Razonar acerca de las distribuciones es más que observar las apariencias, se trata de decodificar las formas mediante el uso de estrategias deliberadas para comprender las distribuciones. Además, los estudiantes tienen que considerar la evidencia para formar una opinión e inferencia a partir de la información contenida en las distribuciones (Friel et al., 2001). Sobre la base de las dificultades que los estudiantes parecían tener para leer e interpretar histogramas, delMas, Garfield y Ooms (2005), se cuestionan si los estudiantes debían ser enseñados a usar solo diagramas de datos y diagramas de caja para representar conjuntos de datos. Concluyeron que había razones importantes para mantener los histogramas en el currículo como una forma de representar las distribuciones de datos, debido a la necesidad de que los estudiantes comprendieran las ideas de área y densidad requeridas para entender las distribuciones teóricas.

2.5 La importancia de entender la inferencia estadística

Existen dificultades para estudiar la inferencia estadística formal en los cursos de bachillerato (Garfield & Ben-Zvi, 2008), esto ha llevado a considerar la Inferencia Estadística Informal (ISI, por sus siglas en inglés) como un concepto general para la enseñanza pre-formal de la estadística (Pratt, 2008). La ISI puede convertirse en un puente entre el análisis exploratorio de datos y la inferencia estadística formal. Si bien el propósito del análisis exploratorio de datos es la exploración sin restricciones de los datos, la búsqueda de patrones interesantes en los datos y los datos disponibles, el propósito de la inferencia estadística es responder a preguntas específicas de una muestra con respecto a la población de donde fue obtenida (Ben-Zvi et al., 2007).

Collins (2003) define a la inferencia estadística como la teoría, los métodos y la práctica de formar juicios sobre los parámetros de una población, a partir del análisis de un evento que involucra un muestreo aleatorio. Moore (2004) argumenta que, la inferencia estadística se mueve más allá de los datos obtenidos de primera mano para sacar conclusiones acerca de un universo más amplio, teniendo en cuenta que la variación está en todas partes y las conclusiones son inciertas. Realizar inferencias de los datos históricos es parte de la vida cotidiana y revisar críticamente los resultados de las inferencias estadísticas de los estudios de investigación es una capacidad importante que todos los adultos, deberían poseer.

La inferencia estadística es el conjunto de métodos y técnicas que permiten inducir, a partir de la información empírica proporcionada por una muestra, cual es el comportamiento de una determinada población con un riesgo de error medible en términos de probabilidad. Los métodos de la inferencia estadística se pueden dividir, en dos: métodos de estimación de parámetros y métodos de contraste de hipótesis. Ambos métodos se basan en el conocimiento teórico de la distribución de probabilidad del estadístico muestral que se utiliza como estimador de un parámetro.

La estimación de parámetros consiste en asignar un valor concreto al parámetro o parámetros que caracterizan la distribu ción de probabilidad de la población. Mientras que los métodos de contraste de hipótesis tienen como objetivo comprobar si un determinado supuesto referido a un parámetro poblacional, o parámetros análogos de dos o más poblaciones, es compatible con la evidencia empírica contenida en la muestra.

El razonamiento acerca de la asociación (o relación) entre dos variables, también denominado razonamiento covariacional, o razonamiento sobre datos bivariados, implica saber cómo analizar e interpretar una relación entre dos variables. La covariación entre eventos o razonamiento sobre datos bivariados, es una base necesaria para inferir una relación causal (Zimmerman, 2005). El tema de covariación, es donde los estudiantes pueden aplicar su conocimiento de inferencia para establecer e interpretar relaciones entre los coeficientes de correlación y las pendientes de regresión.

El razonamiento covariacional también se caracteriza en términos de las actividades cognitivas involucradas en la coordinación de dos cantidades variables. Este tipo de razonamiento puede tomar una forma matemática (por ejemplo, una función lineal), una forma estadística (razonamiento acerca de un diagrama de dispersión) o una forma más cualitativa (por ejemplo, predicciones causales sobre eventos, basadas en asociaciones observadas, como pasar más tiempo estudiando parece conducir a mejores calificaciones de las pruebas, como se describe en la teoría del modelo causal en la psicología).

2.6 Herramientas tecnológicas para la enseñanza de estadística y probabilidad

Los tipos de tecnología utilizados en la estadística y la enseñanza de probabilidad pueden dividirse en varias categorías: software estadístico, software educativo, hojas de cálculo, aplicaciones autónomas, calculadoras gráficas, materiales multimedia y repositorios de datos.

Santos-Trigo (2010a) menciona que, el uso de diferentes herramientas digitales ofrece la posibilidad de examinar tareas matemáticas desde distintas perspectivas que incluyen aproximaciones visuales, gráficas, numéricas y algebraicas. Además, la interacción con tecnologías digitales permite la utilización de imágenes visuales de las ideas matemáticas y la observación de comportamientos o relaciones trascendentes en un problema; organizar y analizar los datos del problema; representar datos en distintos sistemas como el tabular, gráfico y numérico (Santos-Trigo, 2007). El estudiante también puede establecer conexiones entre diversas áreas de las matemáticas como, por ejemplo: álgebra, geometría, probabilidad y estadística.

2.6.1 Software educativo

Diferentes tipos de software estadístico se han desarrollado exclusivamente para ayudar a los estudiantes a aprender probabilidad y estadística. En este estudio se usó el sistema de Geometría dinámica GeoGebra, debido a que es gratuito y a que puede ser descargado para computadoras, tabletas y teléfonos móviles, y con el cual los participantes podían trabajar en el laboratorio de computo de su escuela o en algún dispositivo fuera de la institución educativa. GeoGebra incluye, representaciones geométricas y algebraicas, una hoja de cálculo.

El potencial de la hoja de cálculo de GeoGebra, es que el contenido de las celdas pueden ser objetos geométricos (puntos, rectas, curvas, etcétera), a los que se pueden aplicar operaciones en la forma común en una hoja de cálculo. Por ejemplo, si se tiene una columna de puntos, es posible marcar esa columna y trasladar todos los puntos simultáneamente mediante un cierto vector, obteniendo los puntos resultantes en otra columna. Además, GeoGebra dispone de la herramienta *Cálculo de Probabilidades*, al pulsar sobre esa herramienta se abre una nueva ventana, que ofrece una gran cantidad de posibilidades para el trabajo con las distribuciones de probabilidad (discretas y continuas) comúnmente estudiadas en los cursos de introducción a la estadística. Finalmente, es propio mencionar que GeoGebra ha incorporado una gran cantidad de comandos relacionados con la probabilidad y la estadística, lo que convierten al programa en una herramienta en la enseñanza de la estadística.

2.6.2 Hojas de cálculo

Hojas de cálculo como Excel (http://office.microsoft.com/) están disponibles en computadoras personales; sin embargo, a muchos expertos en estadística no les parece una buena idea el uso de Excel, algunos lo defienden debido a su uso generalizado en la industria y acceso relativamente fácil (Hunt, 1996), quien menciona también que hay que tener cautela al usar Excel como paquete educativo estadístico, ya que resulta complicado es hacer diagramas de caja y bigotes (boxplot).

2.6.3 Aplicaciones autónomas /Archivo en GeoGebra

Una aplicación autónoma, conocida también como applet, es un componente de software que normalmente realiza una función específica y se ejecuta normalmente en un navegador Web, durante la última década, ha habido un crecimiento extraordinario en el desarrollo de applets en línea que pueden ayudar a los estudiantes a explorar conceptos en un entorno visual, interactivo y dinámico. En esta investigación se les proporciono a los alumnos una aplicación construida en GeoGebra, la cual simula el experimento de lanzar dos dados y sumar las caras superiores. Los estudiantes pueden usar y modificar simulaciones de un evento incrustados en el software para explorar conceptos estadísticos, en el caso del presente estudio el comportamiento de la distribución de frecuencias. Con esta aplicación los participantes tendrán la oportunidad explorar, visualizar, e identificar el comportamiento de la distribución de frecuencias si el número de veces que se simula el experimento aumenta de forma considerable.

2.6.4 Internet

La búsqueda de información sobre conceptos estadísticos, en sitios de internet, permite a los alumnos acceder a conceptos, fuentes de información y contenidos matemáticos. Borba, Clarkson, y Gadanidis (2013), argumentan que el hecho de que los estudiantes tengan acceso continuo a internet, refleja cambios como que el conocimiento matemático en su totalidad ya no es sólo propiedad de los profesores y de los libros de texto, ni está limitado por la forma de comunicación en los libros tradicionales. El conocimiento también se presenta en sitios de información de acceso público como Wikipedia y otros que ofrecen contenido matemático textual, multimodal e interactivo. En esta investigación el internet además de ser una herramienta útil en la búsqueda de conceptos estadísticos, también lo es en la indagación de instituciones que ponen a su disposición información estadística que pueda ser consultada y analizada y resuelvan problemas del mundo real.

Capítulo 3. Elementos metodológicos, diseño del estudio y procedimientos

3.1 Introducción

En este capítulo se describe la metodología empleada en el desarrollo de esta investigación, se caracteriza la investigación, la unidad de análisis, la forma de obtención de los datos y se describe la estrategia didáctica empleada durante el trabajo de campo. Además, se muestra el diseño de las actividades que se utilizaron; algunas características de los participantes y los instrumentos de recolección de la información.

3.2 Características de una investigación cualitativa

Hernández, Fernández y Baptista (2014) argumentan que una investigación cualitativa proporciona información detallada del comportamiento de los datos, abundancia interpretativa, contextualización del ambiente o entorno, detalles y experiencias únicas. Además, aporta un punto de vista actual y lógico de los fenómenos, así como flexibilidad, según se encamina por temas significativos de investigación. Los estudios cualitativos pueden incluir preguntas, durante o después de la recolección de datos y el análisis de datos. Estos autores también enuncian que un estudio cualitativo la unidad de análisis, es necesaria para categorizar, codificar e identificar los temas o segmentos de información, documentos u observaciones que se relacionan con las preguntas de investigación en el estudio. Los temas son las ideas y patrones comunes que se observan a medida que se analizan los datos recopilados. En el presente estudio la unidad de análisis, es lo que muestra cada participante al trabajar de manera individual un conjunto de tareas relacionadas con el desarrollo del programa de probabilidad y estadística I correspondiente al Colegio de Ciencias y Humanidades. Con el primer cuestionario y la actividad del lanzamiento de los dados, se busca abarcó las Unidades I y III (Estadística descriptiva y Probabilidad), mientras que con la actividad del ajuste de la línea recta se abordó la Unidad II (Datos bivariados).

3.3 Los participantes

En el estudio participaron estudiantes de quinto semestre de bachillerato, los cuales cursaban la asignatura de Probabilidad y estadística; el grupo estaba conformado por 13 estudiantes cuyas edades oscilaban entre 17 y 19 años. Algunos de los participantes conocían el uso de Excel y GeoGebra; en el caso de las hojas de cálculo las habían empleado en la mayoría de veces para catalogar y ordenar información. La Tabla 1 muestra qué respondieron al preguntarles sobre si habían usado Excel y GeoGebra previamente al desarrollo de las actividades de esta investigación. Durante la implementación de las actividades, el investigador les mostró el uso de las herramientas digitales para despejar dudas con su uso; además de que las actividades contaban con una descripción detallada para que los estudiantes tuvieran una guía del cómo realizarlas.

Nombre	Edad	Uso de Excel	Uso de GeoGebra
Kenya	17	Si	Si
Nisa Danahe	18	Si	Si
Daniela	17	No	Si
Laura Selene	17	No	Si
Iván	19	Si	Si
Victoria	18	Si	No
Adrián	17	Si	No
Adriana	18	No	Si
Brenda	18	Si	No
Fernanda	19	Si	Si
Alfonso	18	Si	No
Aniloreny	18	Si	Si
Ángel	17	No	Si

Tabla 1. Datos de los participantes del estudio

3.4 Sesiones de trabajo y actividades

3.4.1 Sesiones de trabajo

Las sesiones de esta investigación fueron llevadas a cabo en el aula de cómputo de la escuela en donde estudiaban los participantes, se realizaron en total cinco sesiones, cuatro sesiones de 2 horas y una de 1 hora, haciendo un total de 9 horas de trabajo presencial. Cada uno de los estudiantes contó con una computadora para trabajar de manera individual, además los equipos tenían acceso a internet, Excel y GeoGebra, además de un proyector y pantalla, para que el profesor pudiera guiar las sesiones y todos los participantes pudieran seguir las indicaciones dadas, al trabajar las actividades. El investigador contaba con un bloc de notas, computadora y pizarrón.

3.4.2 Diseño de las actividades

En relación a las tareas matemáticas, son el medio por el cual los estudiantes más que seguir y aplicar un conjunto de procedimientos, tienen la oportunidad de involucrarse en un proceso de cuestionamiento, de búsqueda de relaciones y de reflexión conceptual. Para lograr este objetivo, los estudiantes deben utilizar el conocimiento previo y las habilidades adquiridas que les permitan discutir los problemas, representarlos de diferentes formas para que puedan explorarlos, formular relaciones y conjeturas, engancharse en la búsqueda de diversos argumentos para sustentarlas y comunicar los resultados (Santos-Trigo, 2015).

En la primera sesión los alumnos respondieron un cuestionario (Ver Anexo I), con la finalidad de identificaran los conceptos estadísticos con los que iban a trabajar; además de buscar adentrarlos en el tema. Cuando los alumnos desconocían algún concepto, hicieron uso de internet para solventar sus dudas y no existió ningún tipo de limitante al realizar la búsqueda de información, únicamente se les hizo hincapié en que debían anotar la fuente o citar la dirección electrónica de donde obtuvieron su información.

La intención de la actividad llamada el lanzamiento de un par de dados (la actividad se incluye en el Anexo I), es hacer que los alumnos reflexionaran sobre el hecho de que las intuiciones que se tienen al realizar un experimento aleatorio no siempre son correctas, y se buscó que con el uso coordinado de las herramientas digitales lograran la comprensión del comportamiento de la suma de las caras de los dados (evento aleatorio), por parte de los participantes. Los estudiantes generaron y registraron los datos mediante el lanzamiento de un par de dados y sumaron sus caras superiores además usaron una aplicación autónoma que simulaba el experimento. El profesor recolectó los datos de todos los participantes en tablas de frecuencias y los compartió con el grupo, para que después de manera individual fueran analizados mediante gráficos, medidas de tendencia central y dispersión. Además, se buscó que los alumnos, guiados con una serie de preguntas encontraran la relación que existe entre la distribución de frecuencias y la distribución de probabilidad de la suma de las caras de los dados.

La finalidad de la actividad del ajuste de línea recta, fue que los participantes del estudio recapacitaran sobre el hecho de que, mediante el uso de herramientas digitales y procedimientos estadísticos como la regresión lineal y la correlación, se pueden inferir resultados futuros de una colección de datos históricos reales de una característica poblacional, describiendo la naturaleza e intensidad de las relaciones entre esos rubros mencionados y el tiempo.

Los estudiantes realizaron el ajuste de una línea recta de manera visual a una colección de datos, con ayuda de deslizadores que controlaban la pendiente y la ordenada al origen en GeoGebra. Posteriormente el comando de análisis de regresión de dos variables de esa misma herramienta se les solicitó a los jóvenes que determinaran el modelo teórico. Una vez que tuvieron el ajuste de la línea recta, los alumnos realizaron inferencias matemáticas y dieron respuestas a preguntas que estaban dirigidas a ver la interpretación que les daban a los resultados obtenidos.

Los archivos electrónicos generados por el investigador con los que trabajaron los estudiantes, los archivos donde respondieron las actividades los participantes, sus gráficas en Excel y construcciones en GeoGebra se gestionaron en la plataforma Google Classroom. En la primera sesión, el profesor les solicitó a los participantes su correo electrónico, para posteriormente enviarles un documento que incluía las instrucciones que debían seguir para inscribirse en la plataforma (Figura 1).

Figura 1. Pantalla de inicio de Google Classroom

3.4.3 Primera actividad – Cuestionario sobre evento aleatorio

La intención de este cuestionario era que los alumnos centraran la atención en conceptos relacionados con el azar y la probabilidad, las preguntas giraron en relación a dos temas, el primero está vinculado con el concepto que ellos tienen de un evento aleatorio, además de documentar la forma de reaccionar de los participantes al enfrentarse a un evento aleatorio. El segundo tema estaba relacionado con la ludopatía como enfermedad generada por la adicción a los juegos de azar y el impacto que tiene en la sociedad, las preguntas del primer cuestionario se muestran en el Anexo I.

3.4.4 Segunda actividad – Lanzamiento de dados

El problema de la segunda actividad fue encontrar la relación que existe entre la distribución de frecuencias y la distribución de probabilidad de un evento aleatorio. Para lograrlo, se desarrolló una actividad que con la ayuda de herramientas digitales promoviera un acercamiento de los alumnos a los fenómenos aleatorios.

Lo primero que hicieron los participantes fue producir los datos con los que se va a trabajar en la secuencia didáctica, donde los alumnos realizaron 20 veces el lanzamiento de un par de dados distinguibles y sumaron el valor de las caras superiores, después inventaron 20 resultados que pudieran parecer los obtenidos al lanzar los dados, pero sin hacerlo y por último los alumnos hicieron uso de una aplicación autónoma en GeoGebra en la cual se simulaba 20 veces el experimento (Figura 2).

A los alumnos se les proporcionó dados de diferente color, para identificaran que existen pares de dados diferentes que al sumar sus caras superiores el resultado es el mismo, por ejemplo, si se lanza un dado blanco (b) y uno azul (a), las opciones de que las sumas de las caras arrojen un 4, son (1b, 3a), (2b, 2a) y (3b, 1a).

Una vez que cada alumno obtuvo y registró sus resultados de manera individual, el profesor compiló y sumó los datos de todo el grupo para obtener una tabla con las veces (frecuencias) que ocurrieron los valores de las sumas de los dados de los 3 sucesos. Con esa información los alumnos compararon y encontraron similitudes entre los histogramas de frecuencias, diagramas de sectores, de caja y bigotes, medidas de tendencia central y dispersión, usando la hoja de cálculo de Excel y GeoGebra.

Después de que los alumnos registraron los resultados de la simulación del lanzamiento de los dados, modificaron de 20 a 1000 veces el número que se simulaba el lanzamiento de los dados en la aplicación en GeoGebra (Figura 2), y se les cuestionó a cerca del patrón de comportamiento que sigue la distribución de frecuencias. Los resultados obtenidos por los alumnos con estos acercamientos se buscaron que fueran conectados con el concepto de distribución de probabilidad, cuyo significado debe ser buscado por los alumnos haciendo uso de la internet. Esta actividad se muestra en el Anexo I.

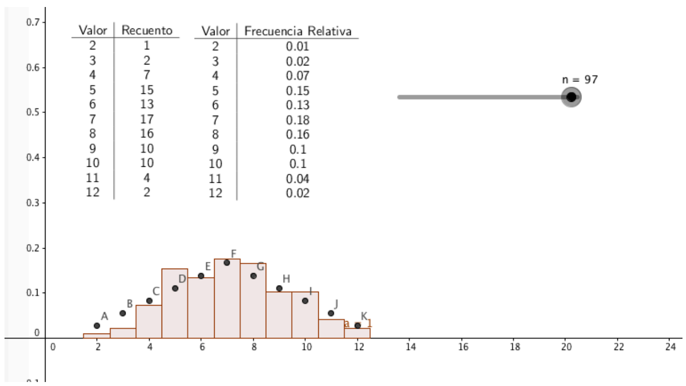


Figura 2. Modelo dinámico construido en GeoGebra

3.4.5 Tercera actividad – la población joven por mediante un ajuste de línea recta

El problema de la tercera actividad fue encontrar el ajuste de regresión lineal de una colección de datos históricos, haciendo uso de una herramienta digital, con la finalidad de inferir resultados futuros.

Previo a iniciar la tercera actividad, el investigador cuestionó a los alumnos, ¿si conocían alguna dependencia o institución que posea información (fidedigna) de la población de México? Esta pregunta tiene la intención de sensibilizar y orientar a los alumnos a que sean capaces de buscar información real y con ella poder emplear los conceptos y conocimientos adquiridos en la escuela.

Posteriormente el profesor les proporcionó un archivo de Excel el cual contenía una tabla con el histórico de la población joven de los países miembros de la Organización para la Cooperación y el Desarrollo Económico (OCDE), esa información se presenta por país, año y la proporción que representan de la población total del país respectivo. Esos datos se gestionaron a través de la plataforma Google Classroom y fueron obtenidos por el investigador del portal de internet de OCDE, ver Figura 3.

https://data.oecd.org/pop/young-population.htm#indicator-chart.

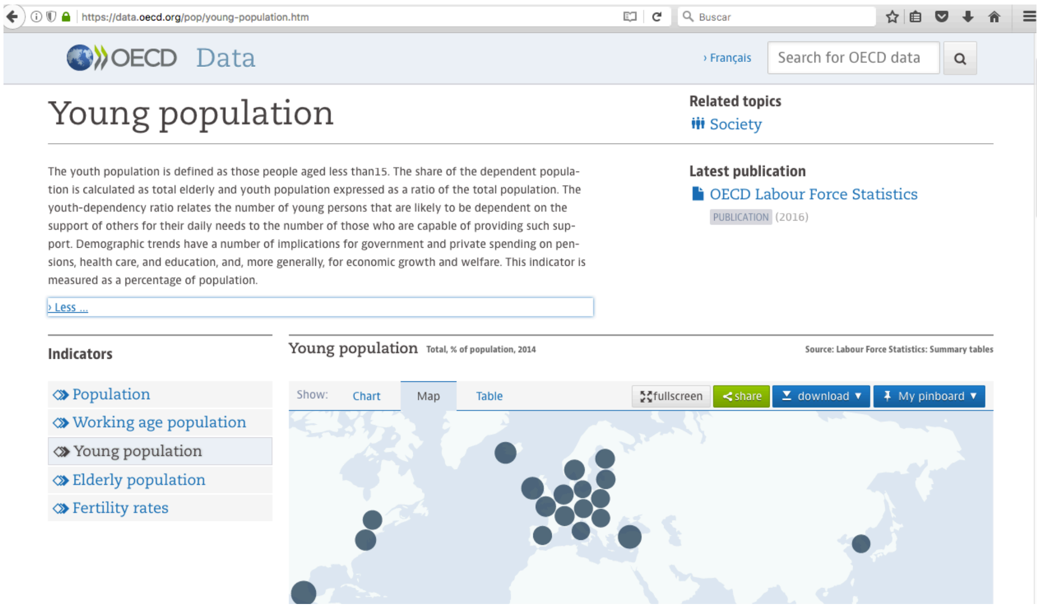


Figura 3. Página de la OCDE con la información de la población joven

Los alumnos del archivo de Excel, únicamente hicieron uso de los de la población joven de México la cual aparece desde 1970 hasta 2014. Cabe mencionar que la OCDE define a la población joven como aquellas personas menores de 15 años y este indicador se mide como un porcentaje de la población. OCDE (2016), Young population (indicator). La actividad completa se incluye en el Anexo I.

Previo a la realización de esta fase existió la necesidad de mostrarles a los participantes el uso del comando *deslizador* en GeoGebra. El deslizador es la herramienta que les permitió modificar los valores de los parámetros de la pendiente m y la ordenada al origen b, con los cuales se buscó realizar el ajuste de la línea recta (ver Figura 4) de forma visual. Cuando se crea un deslizador es necesario definir los valores máximo y mínimo del rango, además el valor del incremento al moverlo, y estas fueron situaciones que los alumnos tuvieron que solventar al responder la actividad.

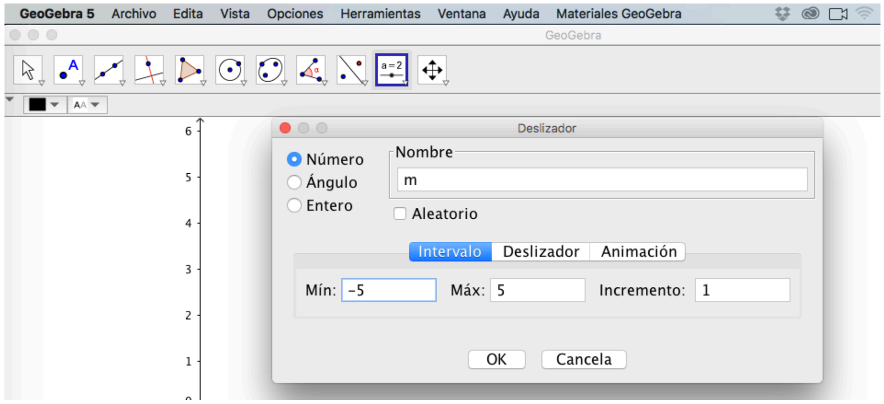


Figura 4. Elementos del deslizador en GeoGebra

Una vez sea solventada esta dificultad se les solicitó a los alumnos que graficaran con puntos los valores de la tabla de Excel, las abscisas fueron la población joven y las ordenadas el respectivo año (diagrama de dispersión) y sobre la misma, trazaron una línea recta la cual fueron ajustando lo mejor posible a dicho diagrama; para ello, los jóvenes diseñaron un par de deslizadores que controlaban los parámetros de la pendiente y la ordenada al origen de la mencionada recta, fue necesario que los participantes visualizaran el comportamiento de los datos y así encontraron el patrón que seguían. Este fue un primer acercamiento que favoreció el acceso al concepto de ajuste de regresión lineal.

Una vez que los alumnos realizaron el ajuste de la línea recta de manera visual, seleccionaron todos los datos nuevamente y con el comando de análisis de regresión de dos variables de GeoGebra determinaron de forma analítica la línea recta ajustada, haciendo uso del comando de análisis de regresión de dos variables. Después compararon ambos ajustes e identificaron similitudes y diferencias, además ya que contaban con el modelo analítico realizaron inferencias de resultados futuros, calcularon la población joven de México en el año 2020, y determinaron el año en el que este indicador poblacional aproximadamente seria cero, además de verificaron la intensidad de la relación entre la población joven y el tiempo, haciendo uso del concepto de coeficiente de correlación, el cual se les pidió que investigaran en internet. Al finalizar la actividad, los participantes respondieron a la pregunta, ¿qué acciones o medidas en términos de políticas públicas podrías sugerir que se tomen en cuenta con base a los resultados obtenidos?

3.5 Instrumentos de recolección de información

Los archivos entregados por parte del investigador hacia los participantes, así como los archivos generados por los mismos, fueron gestionados a través de la plataforma Google Classroom. El que suscribe creó una clase llamada "Probabilidad y estadística 1" (ver Figura 5), en donde los alumnos fueron dados de alta en la misma mediante su cuenta de correo electrónico y una vez matriculados tenían acceso a los archivos, así como la opción de entregar los archivos que ellos mismos generaron.

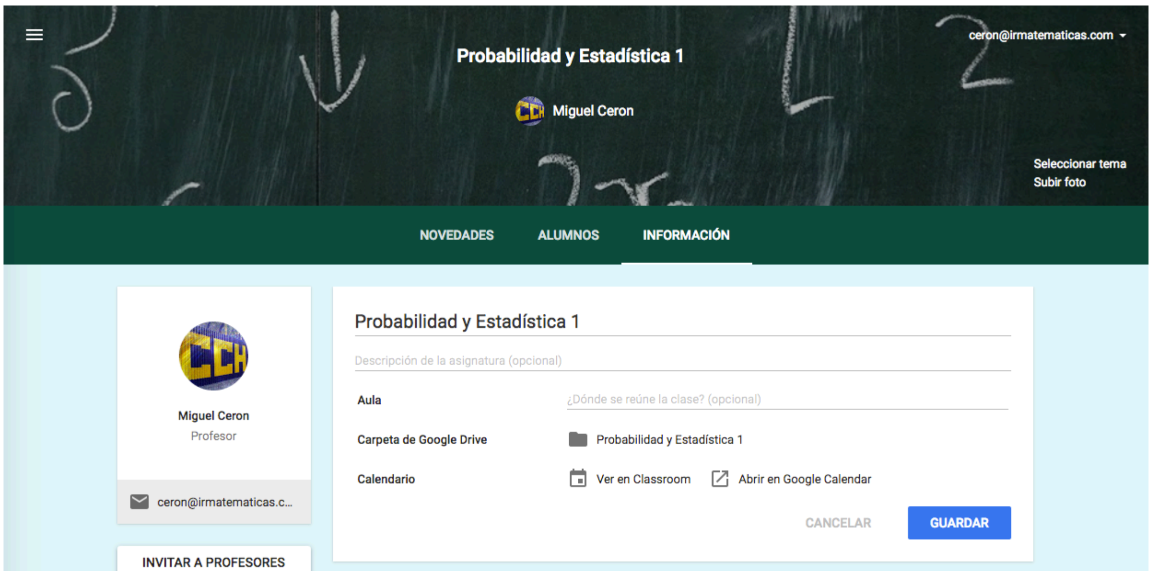


Figura 5. Pantalla de inicio en Google Classroom de la clase "Probabilidad y estadística 1"

Archivos de Excel: para la actividad del lanzamiento de los dados los alumnos entregaron, los archivos que contienen las gráficas realizadas por ellos y que fueron hechas con la información de los dados lanzados, inventados y simulados; además del gráfico comparativo de los datos antes mencionados. Mientras que para la segunda actividad el investigador a través de Google Classroom puso a disposición de los participantes un archivo en Excel el cual contenía la información de la población juvenil de México desde 1970 hasta 2014.

Archivos de GeoGebra: en la actividad del lanzamiento de los dados, los participantes realizaron un comparativo de los datos de los dados lanzados, inventados y simulados, haciendo uso del comando "análisis multivariable", de esta forma obtuvieron los diagramas de caja y bigotes de dichos datos, así como las medidas de tendencia central y dispersión de los mismos, el archivo generado por ellos les fue solicitado y entregado a través de la plataforma de gestión de archivos ya mencionada. Para la actividad del ajuste de recta los participantes trabajaron en GeoGebra, exportando a la vista gráfica de la herramienta los datos de un archivo de Excel, para que en un principio construyeran un par de deslizadores que controlaban los parámetros de la pendiente y la ordenada al origen y así pudieran realizar un ajuste visual, para después usar el comando análisis de regresión de dos variables de la herramienta.

Archivos en Word: las actividades, así como los cuestionarios que los alumnos respondieron fueron entregadas en este formato y una vez que los participantes iban contestándolas, estas eran entregadas a través del mismo gestor de archivos Google Classroom.

Notas de campo: estas son las anotaciones que el investigador realizó como evidencia escrita y que fueron tomadas en el proceso de la recolección de la información durante las cinco sesiones. En estas notas se encuentran las dudas de los participantes con el uso de las herramientas tecnológicas, comentarios y preguntas de los alumnos durante la realización de las actividades. Estas anotaciones también sirvieron para complementar la información que los alumnos entregaron a través de los archivos de Word, Excel y GeoGebra, para su posterior análisis.

3.6 Estudio estadístico de las actividades

Las actividades fueron diseñadas con base al curriculum del curso de probabilidad y estadística 1, perteneciente al Colegio de Ciencias y Humanidades (CCH), que es un bachillerato de la Universidad Autónoma de México (UNAM). El temario considerado consta de 3 unidades. La primera unidad es estadística descriptiva, la segunda es datos bivariados y la última es probabilidad.

La primera y segunda actividad que son el cuestionario sobre el evento aleatorio y el lanzamiento de dados, fueron pensados en abarcar las unidades de estadística descriptiva y probabilidad. Mientras que la tercera actividad, llamada ajuste de una línea recta está diseñada para abarcar la unidad de datos bivariados.

3.6.1 Estudio del lanzamiento de los dados

Existen diferentes corrientes de probabilidad (frecuentista, clásica, subjetivista, axiomática y bayesiana), las cuales se aplican para asignar un valor numérico a la posibilidad de la ocurrencia de algún suceso probabilístico el cual ocurre bajo cierto grado de incertidumbre.

En la *corriente frecuentista* se asigna un valor de probabilidad a un evento *E*, a partir de lo que se considera que ocurrirá. Su definición o interpretación de la probabilidad está basada, en la frecuencia relativa[5] con la que se obtendrá *E*, si el experimento se repite una gran cantidad de veces en condiciones similares (no idénticas, puesto que en este caso el proceso no sería aleatorio).

En la *corriente clásica* (a priori) se consideran espacios muestrales uniformes, es decir que se asigna probabilidades a eventos, basándose en resultados equiprobables (igualmente verosímiles). Esto es, que se asignan la misma probabilidad a cada punto del espacio muestral (1/n, en donde n es la cantidad de elementos del espacio muestral), posteriormente para obtener la probabilidad de la ocurrencia de un evento *E*, se suma la cantidad de elementos de *E*, y se multiplica por la probabilidad de un elemento del espacio muestral (1/n). Es importante mencionar que la probabilidad de los puntos muestrales se establece a priori, es decir antes de la realización de cualquier experimento.

En una primera instancia los alumnos deben de hacer uso de la estadística descriptiva, para recopilar, graficar y describir el comportamiento de los datos obtenidos al lanzar, inventar y simular los datos obtenidos al sumar las caras superiores de los dados. Los participantes deben tener en cuenta que existen diferentes pares ordenados que dan un mismo resultado. Por ejemplo, el 7, lo generan diferentes pares 6+1, 5+2, 4+3, 3+4, 2+5 y 1+6, mientras que para obtener el 12, solamente existe uno solo, es decir 6+6.

En consecuencia, al simular el lanzamiento de los dados de forma aleatoria y posteriormente sumarlos un número considerable de veces, la frecuencia de obtener el 7 es mayor a la frecuencia de obtener un 12. De forma análoga, sucede con los demás números. Precisamente, poder reproducir un número considerable de veces el experimento, se debe emplear la aplicación que lo simula, en la cual se registran las frecuencias absolutas, relativas y se traza la gráfica. La obtención de la distribución de frecuencias es un acercamiento al concepto de distribución de probabilidades (Ver figura 6).

[5] La frecuencia relativa de un suceso es igual al cociente de las veces en que ocurre el suceso entre el total de veces que se repita el experimento

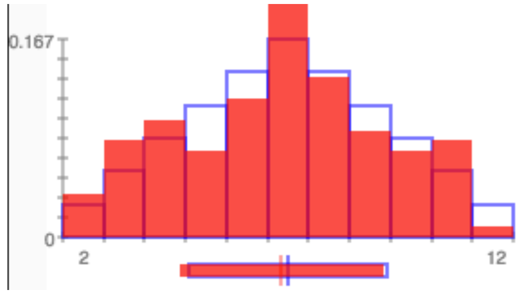


Figura 6. Comparativo de la distribución de frecuencias y la distribución de probabilidades de lanzamiento de un par de dados

La distribución de probabilidades se obtiene al determinar las probabilidades de manera clásica, tomando en cuenta que el espacio muestral del lanzamiento de los dos dados es 36. Los posibles valores al sumar las caras superiores de los dados son 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 y 12 (ver Figura 7). Y el número de pares que los generan son 1, 2, 3, 4, 5, 6, 5, 4, 3, 2 y 1, respectivamente.

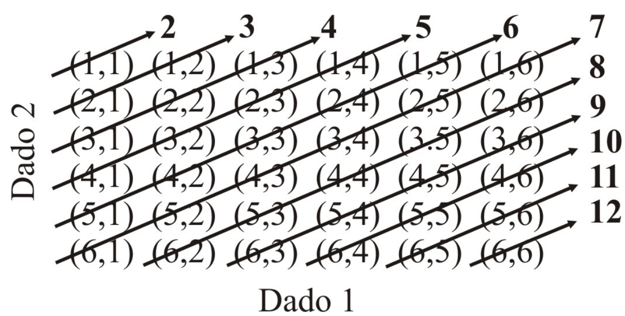


Figura 7. Espacio muestral del lanzamiento de un par de dados

En consecuencia, la distribución de probabilidades del lanzamiento y suma de las caras superiores de dos dados (ver Figura 8), es:

x	2	3	4	5	6	7	8	9	10	11	12
p(x)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Figura 8. Distribución de probabilidad para el lanzamiento de un par de dados

Este es el resultado que se espera que los alumnos obtengan al final de la primera y segunda actividad.

3.6.2 Estudio de la población joven por ajuste de una línea recta

El modelo de regresión lineal y el análisis de correlación son herramientas importantes en los campos de la ciencia cuyos resultados provengan de variables cuantitativas, tanto discretas como continuas. Al estudiar un fenómeno, mínimo se examinan dos variables, cuyo propósito, por lo general consiste en usar una para pronosticar la otra. Así, la mayor parte de los estudios de regresión se inician con el deseo de examinar el valor cambiante de una variable, que en el análisis de regresión se llama variable dependiente.

El símbolo elegido para la variable dependiente es y. Se identifica una segunda variable que se piensa está asociada a y y se le llama variable independiente o predictora, su símbolo es x. Es necesario conocer el comportamiento de la dependencia entre las variables definidas y esto se logra trazando un diagrama de dispersión, el cual es una representación gráfica de dos variables cuantitativas que se analizan de manera simultánea. El objetivo principal del análisis de regresión es predecir el valor de una variable (la variable dependiente) dado el valor de una variable asociada (variable independiente).

La característica de los diagramas de dispersión es que los datos se presentan en forma de puntos, sin estar unidos por segmentos de recta. La escala del eje X contiene el rango de los valores necesarios para la variable x, y el eje Y también tiene una escala adecuada para los valores de y. Los pares de datos se representan gráficamente en un sistema de dos dimensiones. Una ventaja del diagrama de dispersión es que permite ver la relación entre las dos variables de interés. Y al trazarlo se puede descubrir la presencia de relaciones entre las variables y su tipo e identificar si la relación es lineal o no.

El término de análisis de regresión simple indica que la variable dependiente se predice sobre la base de una sola variable independiente, el tipo más sencillo de curva de aproximación en un modelo es la línea recta y cuando se examina la relación de dos variables, por lo general, se hace con el propósito de usar una para pronosticar la otra.

En el caso de la tercera actividad propuesta, a los alumnos se les proporciona un archivo en formato Excel el cual contiene información de la población joven de México y en donde los valores de los años son representados por la variable independiente x, mientras que el valor de la proporción de población joven está representado por la variable *y*. Y al ser graficados y analizados, se observa un descenso de forma constante y sigue un comportamiento casi lineal. Este fenómeno se observa debido a que como lo menciona el Consejo Nacional de Población (CONAPO, por sus siglas en español, 2013), se confirman el descenso en la fecundidad, la mortalidad general y la infantil, así como el aumento en la esperanza de vida de la población en general, el rezago en el bienestar y condiciones de vida de determinados sectores de la población son factores para que exista una transición de forma lenta, pero constante.

La fecundidad en México continúa en franco descenso y en el último año del primer cuarto de este siglo estará cada vez más cerca del nivel del reemplazo (que es de 2.1 hijos por mujer). En la última década del siglo pasado ocurrió una disminución considerable en el número promedio de hijos por mujer al pasar de 3.36 a 2.65. Al presente, la tasa global de fecundidad es de 2.22 hijos por mujer. A finales de los años setenta comienza a disminuir la fecundidad en México, debido principalmente a la instrumentación de acciones en materia de planificación familiar. En los primeros veinte años (entre 1970 y 1990) se apreció una acelerada reducción de la Tasa Global de Fecundidad (TGF), por ejemplo, de 1970 a 1980 descendió en casi dos hijos y de 1980 a 1990, en poco más de un hijo.

Así, en 1990, el número de hijos que tuvieron las mujeres fue la mitad de los que tenían en 1970; así mismo, el descenso fue continuo, pero a un ritmo menos acelerado, es decir, tuvieron que pasar veinte años (1990 a 2010) para reducir la TGF en alrededor de un hijo y se espera que en las próximas dos décadas (2010 a 2030) se logre estar alrededor del nivel de reemplazo generacional (2.1 hijos por mujer).

La CONAPO (2013), menciona que la reducción de la fecundidad ha sido favorecida por la implementación de programas gubernamentales de planificación familiar que promovieron el uso de métodos anticonceptivos; en un principio, sólo las mujeres residentes de zonas urbanas tuvieron acceso a estos beneficios, pero con el paso del tiempo se ha buscado que la cobertura se extienda al resto de las mujeres del país.

El uso de métodos anticonceptivos entre mujeres unidas en edad fértil a nivel nacional se ha incrementado, en 1976 fue de 30.2 por ciento, en 1987, de 52.7, en 1997, de 68.5, y en 2009, de 72.5 por ciento. Estos datos muestran que la mayor cobertura del uso de métodos anticonceptivos sucedió en las dos primeras décadas, y dicho indicador todavía continúa en ascenso, lo que reafirma la influencia que han tenido las políticas públicas en la disminución de la fecundidad.

Las entidades que en 2010 tuvieron la TGF más alta fueron Chiapas, Guerrero y Oaxaca, que por lo general se caracterizan por tener grados de marginación muy alto o alto, tienen un mayor porcentaje de población rural, que habla lengua indígena o que vive en localidades de difícil acceso; además, su prevalencia anticonceptiva fue baja. En contraste, el Distrito Federal y estados como Baja California Sur y Baja California registraron las tasas de fecundidad más bajas, y cuentan con todos los servicios necesarios para que la población tenga información y acceso a la planificación familiar.

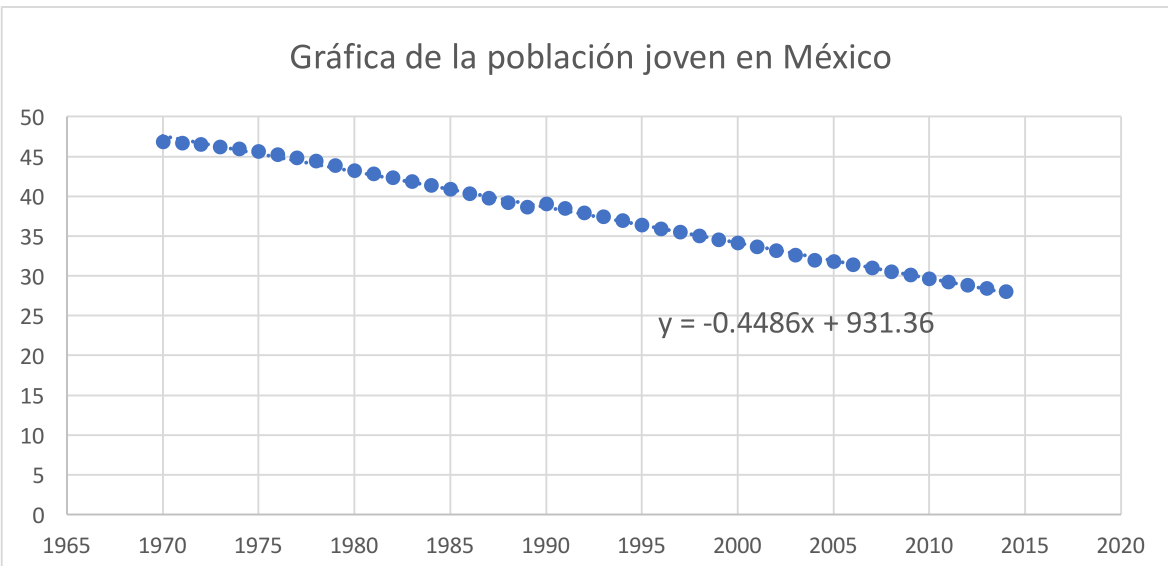


Figura 8. Espacio muestral del lanzamiento de un par de dados

Al realizar el ajuste de la línea recta con los datos de la población joven, proporcionados por la OCDE se obtiene el modelo $y = -0.45x + 931.36$, donde la ordenada al origen es $y = 931.36$, el cual es valor esperado de y si $x = 0$, (ver Figura 8). De la ecuación de regresión, se puede observar que por un aumento de una unidad en x implica que el valor de y disminuya en promedio 0.45 la proporción de la población joven. En términos prácticos, la ecuación de regresión sugiere que para cada año que transcurra, se puede esperar un promedio de 0.45 de disminución en el índice de proporción de población joven. Lo cual concuerda con la información proporcionada por la CONAPO, ya que algunos de los principales factores que impactan el índice de población joven en México, el control de la tasa de fecundidad mediante la proliferación de los anticonceptivos, además del aumento en el costo de la vida, el encarecimiento de la vivienda y algunos otros rubros que hacen que las familias cada vez tengan menos hijos.

Es de llamar la atención que, en el caso de los datos de la población joven de México, desde 1970 hasta 2014, siga un comportamiento lineal casi perfecto, con una pendiente negativa. Esto no significa que se tengan que cumplir con ese comportamiento los años venideros, a menos que las condiciones demográficas y los factores que impactan el índice de población joven, se mantengan para generar dicho comportamiento lineal.

3.7 Análisis de datos

Al realizar el análisis de los datos se proporcionó una estructura que permitió reducir los datos, organizarlos y verificar los resultados. Los datos que se obtienen en una investigación cualitativa son muy diversos, por ejemplo, visuales, textos escritos, expresiones verbales, además de las narraciones del investigador, por eso los objetivos principales del análisis cualitativo son explorar los datos, organizarlos en unidades y categorías, describir las experiencias de los participantes, descubrir los conceptos y patrones presentes en los datos para interpretarlos y explicarlos en función del planteamiento del problema, comprender el contexto y asociar los resultados con el conocimiento disponible (Hernández, Fernández y Baptista, 2014).

Los datos se analizaron comparando, la información recolectada y consistió en tres fases: reducción de datos, organización y despliegue de datos, y elaboración de las conclusiones. En la fase de reducción, se seleccionaron, simplificaron y resumieron los datos recolectados mediante la identificación de las regularidades, similitudes, patrones, diferencias y errores, siguiendo como marco metodológico, la resolución de problemas propuesto por Santos-Trigo y Camacho-Machin (2009).

Para la reducción de datos, se emplearán los archivos de Word con las respuestas que dieron los estudiantes a los cuestionarios y a las actividades matemáticas, además de incluir los archivos entregados en Excel y en GeoGebra, donde los alumnos realicen las gráficas, así como todo el trabajo que se obtenga del desarrollo de las actividades.

La fase de organización y el despliegue de datos, consistió en la presentación organizada de los datos, de modo que sea posible observar las regularidades, diferencias, errores y patrones. La presentación de los resultados obtenidos en las actividades, se organizaron de acuerdo a los episodios propuestos por Santos-Trigo y Camacho-Machin (2009).

En la última fase se la formularon conjeturas y verificaron conclusiones, considerando las regularidades y patrones identificados durante el análisis de datos, contrastándolas con las ideas expuestas en el marco conceptual con el fin de dar respuesta a las preguntas de la investigación planteadas y realizar algunas reflexiones sobre el problema de investigación propuesto en el trabajo.

Sesión	Actividad	Trabajo realizado
Sesión 1	Alta de los participantes a la plataforma Google Classroom. Resolución de la actividad 1	• Se les solicitó a los participantes su cuenta de correo, con el fin de darles de alta en la plataforma Google Classroom y puedan gestionar los archivos de esta investigación. • Se les pidió a los participantes realizaran la primera actividad que es el cuestionario sobre el evento aleatorio, los jóvenes tuvieron acceso a internet, con el fin de que, si desconocían el concepto, pudieran buscar información relacionada con el tópico y así contestaran todas las preguntas del cuestionario.
Sesión 2	Revisión de la actividad 1 Resolución de la actividad 2	• De manera grupal se revisan las dudas e inquietudes relacionadas con las preguntas del cuestionario. • Se inicia la segunda actividad llamada el lanzamiento de los dados. Los alumnos generaron la información de los dados, primero de manera individual, después el investigador recolectó la información de todos los participantes y la compartió con los estudiantes para procedieran a la graficación de los datos resultantes.
Sesión 3	Conclusión de la actividad 2	• En esta sesión se continuo con la observación y comparación de las gráficas de barras y el diagrama de caja y bigotes de los datos obtenidos. Además, los participantes calcularon y registraron las medidas de tendencia central y de dispersión. • Para finalizar la actividad, los estudiantes investigaron el concepto de distribución de probabilidad, además se les pidió que calcularan esta distribución para la suma de las caras de un par de dados. Una vez que tuvieron la distribución probabilidad, la compararon con la distribución de frecuencias al simular el lanzamiento de los dados con la aplicación hecha en GeoGebra.
Sesión 4	Resolución de la actividad 3	• Se exhortó a los alumnos a que investigaran sitios de internet de entidades de gobierno o diversas instituciones donde se pudiera acceder a información demográfica de México. • Posteriormente el investigador les proporcionó a los alumnos la información de la población joven de México, obtenida del portal de la OCDE. • Una vez que los alumnos contaban con la información antes señalada, haciendo uso de GeoGebra, graficaron los valores del año y el valor de la población joven en un diagrama de dispersión. Una vez obtuvieron la gráfica de dispersión, se les pidió que identificaran la forma geométrica que aparentan tener los datos.

		• Al descubrir que los datos aparentan ser una línea recta, los participantes crearon un par de deslizadores que controlaban el movimiento de los parámetros de la pendiente y ordenada al origen, los cuales fueron usados para realizar el ajuste de una línea recta de forma visual.
Sesión 5	Conclusión de la actividad 2	• En la última sesión los alumnos realizaron el ajuste de línea recta de los datos de la población joven proporcionada por el investigador, empleando la herramienta tecnológica. • Una vez que los participantes obtuvieron el ajuste de la línea recta, se realizaron inferencias matemáticas. Y con esa información dieron respuesta a una serie de interrogantes que tienen que ver con las posibles acciones a considerar por parte de gobierno y así ellos pudieran hacer latente el uso de la estadística como una herramienta que puede ser usada en la inferencia de resultados futuros y la toma de decisiones.

Capítulo 4. Análisis de Datos

En este apartado se muestran y analizan los resultados obtenidos durante las etapas del estudio: aplicación de los cuestionarios iniciales y la resolución de actividades de trabajo. En las actividades propuestas, los alumnos hicieron uso de GeoGebra y hojas de cálculo para realizar representaciones gráficas, cálculos numéricos, determinar medidas de tendencia central y de dispersión, así como hacer inferencias.

Se describen las estrategias de cada alumno al resolver cada una de las tareas con el fin de determinar el nivel que se encuentra en cada proceso de pensamiento: descripción, organización, representación, análisis e interpretación de datos e inferencia de resultados.

Previo a la realización de la primera actividad (llamada el lanzamiento de dados), se presentó un cuestionario el cual tenía como finalidad identificar los conocimientos y conceptos que poseían los participantes del estudio, relacionados con el tema de evento aleatorio.

Los resultados obtenidos en las actividades, están organizados de acuerdo a los episodios propuestos por Santos-Trigo y Camacho-Machin (2009).

4.1 Primera actividad – Análisis del cuestionario sobre el evento aleatorio

En la primera sesión de trabajo, se les pidió a los participantes que contestaran un cuestionario. Los participantes tuvieron acceso a internet, con el fin de que, si desconocían algún concepto, tuvieran oportunidad de buscar información relacionada con el tópico y pudieran contestar todas las preguntas. Se hizo énfasis que debían citar las fuentes consultadas y explicar el concepto con sus propias palabras. También se les cuestionó acerca de sus intuiciones al enfrentarse a un evento aleatorio y cómo está relacionada la ludopatía con los conocimientos que una persona posee acerca de la probabilidad.

Al cuestionamiento, ¿qué entiendes por un evento aleatorio?

El dominio del conocimiento que los alumnos poseen del tema de evento aleatorio está orientado en dos vertientes, por un lado 8 alumnos concordaron en que un evento aleatorio está relacionado de alguna forma con el azar o la suerte, mientras que los otros 5 alumnos reportan que, son situaciones donde existen diferentes resultados, ya que en un experimento aleatorio no se puede predecir o reproducir el resultado exacto de cada experiencia particular y puede presentar resultados diferentes debido a que este se rige por el azar; esto es un avance en el sentido de que se trata de una concepción superficial del concepto de distribución de probabilidad (Ver Figura 1).

Respuesta	**Alumnos**
Un evento aleatorio está relacionado de alguna forma con el azar	Kenya, Nisa Danahe, Daniela, Laura Selene, Iván, Victoria, Adrián y Fernanda
En un evento aleatorio existen diferentes resultados	Adriana, Brenda, Alfonso, Aniloreny y Ángel

Figura 1. Respuesta de los alumnos a la pregunta, ¿qué entiendes por evento aleatorio?

Al solicitarle a los participantes ejemplificar eventos aleatorios y fenómenos aleatorios en su vida. Todos los participantes apelaron a los recursos básicos y conocimientos previos que poseen mencionando el lanzamiento de una moneda, el lanzamiento de dos dados, los juegos de lotería, la repartición de cartas y la extracción de bolas de una urna. En el caso de los ejemplos en su vida cotidiana, mencionaron fenómenos meteorológicos como la lluvia, el frío, el calor o el viento, así como situaciones de tránsito vehicular o el tiempo que tarda la llegada del metro.

Al cuestionar a los participantes, ¿de qué forma reaccionas al enfrentarte a un experimento aleatorio?

Nuevamente los participantes hacen uso de los conocimientos previos que tienen y refieren a los ejemplos expuestos anteriormente de manera que Daniela y Alfonso respondieron que lo único que los puede ayudar es su intuición. Los demás participantes reportan que el azar no obedece a ninguna ley o patrón y mencionan que al encarar un experimento únicamente piensan en la suerte. Schoenfeld (1985), menciona que una parte de los recursos que posee un individuo es el conocimiento informal e intuitivo, el cuál es la forma de interpretar conocimiento nuevo por parte de esa persona, en el caso de los alumnos mencionan que el azar no sigue ningún comportamiento o patrón, es necesario confrontar esta idea.

Las respuestas que los alumnos dieron a la pregunta, ¿qué importancia tiene el educar a los alumnos en temas que impliquen fenómenos aleatorios?, los alumnos hacen uso de su sistema de creencias coincidiendo en que es relevante aprender el comportamiento de los fenómenos aleatorios y que puede ayudarlos a tomar decisiones correctas o favorables. Incluso, Laura Selene y Adriana reportan que la experimentación y el uso de la simulación puede ayudarlos a tomar decisiones convenientes (Ver Figura 2).

Laura Selene:

Porque con la experimentación y la simulación de eventos aleatorios nos ayudará a superar las dificultades y obstáculos que podemos tener sobre el desarrollo de la intuición del azar, esta se desarrolla después de un proceso de enseñanza.

Adriana:

La importancia es que puedes ayudar a muchas personas en la toma de decisiones, ya que al saber sacar los posibles resultados de algo puede saber que le conviene y que no a la hora de tomar la decisión.

Figura 2. Evidencias de Laura Selene y Adriana

A la interrogante de si ¿conoces qué es la ludopatía?

Todos los estudiantes recurrieron a internet para consultar el término, reportando la definición encontrada, la fuente y una explicación con sus propias palabras como se muestra en las respuestas de Kenya y Adriana (Ver Figura 3).

Kenya:

Definición de:
https://archivos.juridicas.unam.mx/www/bjv/libros/6/2803/13.pdf
Paráfrasis:

La ludopatía es una enfermedad adictiva, se me ocurre algo parecido al T.O.C caracterizada por tener una adicción a los de juegos de azar, se caracteriza por tener un deseo incontrolable de jugar.

Adriana:

La ludopatía se caracteriza fundamentalmente porque existe una dificultad para controlar los impulsos, y que en cierto sentido tiende a manifestarse en practicar, de manera compulsiva, uno o más juegos de azar.
https://es.wikipedia.org/wiki/Ludopat%C3%ADa
Enfermedad psicológica que lleva a la adicción de los juegos de azar.

Figura 3. Evidencias de Kenya y Adriana

La importancia que aporta el internet como herramienta de búsqueda de información, queda evidenciada en las respuestas que ofrecen los alumnos a la respuesta de si conocían qué es la ludopatía (Figura 3), en donde después de realizar la búsqueda de concepto, citaron la fuente o de donde obtuvieron la información, además de ofrecer una explicación, exponiendo así que no copiaron y pegaron la información.

Las réplicas que los participantes reportan a la pregunta, ¿qué tan importante es que la población en México reconozca a la ludopatía como una enfermedad?

El objetivo de la pregunta fue analizar si los alumnos habían entendido el término de ludopatía e identificar el sistema de creencias de los participantes. Todos los alumnos argumentan la importancia de reconocer que la ludopatía sea considerada una enfermedad, como lo escribe Ángel (Ver Figura 4).

Ángel:

Yo considero que si es muy importante que se hable sobre el tema y se informe porque muchas personas no conocen el término y quizá alguien esté pasando por esa situación y así podrían ser tratadas a tiempo porque en lo personal he conocido personas que por el vicio de juegos de azar llegan a apostar cosas importantes como su casa e incluso pierden a su familia.

Figura 4. Evidencia de Ángel

La última pregunta del cuestionario fue, ¿crees que los conocimientos acerca de probabilidad y estadística influyan, para que una persona sea menos propensa a ser ludópata?

Respuesta	**Alumnos**
Los conocimientos que una persona posea de probabilidad y estadística no influyen para ser propensos a ser ludópatas	Nisa Danahe, Alfonso, y Fernanda
El tener conocimientos de probabilidad puede ayudar a las personas a ser menos propensos a ser ludópatas	Kenya, Daniela, Laura Selene, Iván, Victoria, Adrián, Adriana, Brenda, Aniloreny y Ángel

Figura 5. Respuesta de los alumnos a la pregunta, ¿crees que los conocimientos acerca de probabilidad y estadística influyen para que una persona sea menos propensa a ser ludópata?

En las respuestas de Nisa Danahe, Alfonso y Fernanda, se evidencian el uso de la heurística de la disponibilidad argumentando que una vez que gana una persona por un golpe de suerte, seguirá ganando en los juegos subsiguientes, de hecho, este tipo de sesgos los explotan los juegos de lotería y sorteos, debido a que, si las personas comprendieran verdaderamente las probabilidades que tienen de ganar, seguramente muy poca gente se animaría a comprarlos (Ver Figura 5).

El resto de los alumnos consideran que el tener conocimientos de probabilidad puede ayudar a ser menos propensos a padecer la ludopatía. En las respuestas se identifican que los alumnos reconocen a los conocimientos de probabilidad y estadística, como herramientas que resultan útiles en la comprensión del comportamiento de los fenómenos aleatorios; como lo mencionan Laura Selene y Ángel (Ver Figura 6).

Laura Selene:

> Yo pienso que el conocimiento de la probabilidad y la estadística son de ayuda para que seamos menos propensos a este problema, ya que nos hace ver estos juegos de azar como una parte fundamental de las matemáticas y de la vida en sí, porque existen fenómenos aleatorios en nuestra vida cotidiana y la información acerca de estos juegos nos ayudara a prevenirla.

Ángel:

> Si porque así podría tener noción de que tanta probabilidad hay de que gane o pierda antes de apostar o estaría consiente de las consecuencias que le podría traer.

Figura 6. Evidencias de Laura Selene y Ángel

Estas preguntas sirvieron para que el investigador conociera los recursos y creencias que tenían los participantes al enfrentarse a un evento aleatorio.

4.2 Segunda actividad – Lanzamiento de dados

La primera parte de la actividad consistió en que cada participante: 1) realizara 20 veces el lanzamiento de dos dados distinguibles de color diferente y registraran el valor de la suma de las dos caras superiores; 2) inventara 20 resultados que puedan parecer los obtenidos al lanzar los dados y 3) utilizaran una aplicación en GeoGebra (Figura 7) la cual simula 20 veces el experimento de lanzar dos dados y obtener la suma de sus caras superiores. En la simulación se muestran, a) la tabla con el número de veces que aparecen los valores de la suma de los dados (frecuencia absoluta), b) la tabla con el porcentaje que le corresponde a cada valor de la suma de los dados (frecuencia relativa), c) la gráfica de barras de las frecuencias relativas, d) las marcas de clase para una distribución de probabilidad y e) el deslizador que controla el número de veces que se simula el lanzamiento.

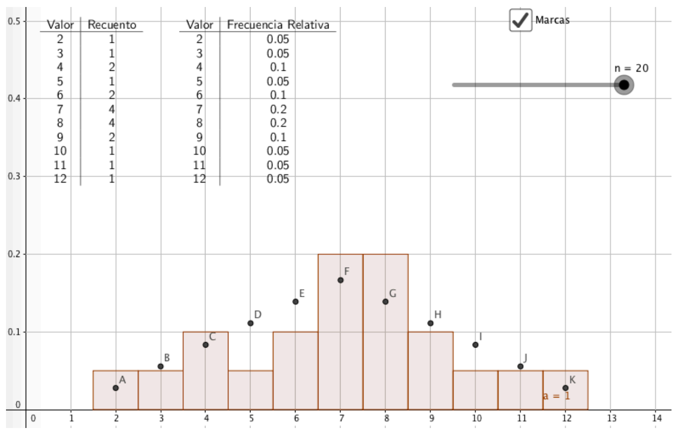


Figura 7. Aplicación que simula el lanzamiento de los dados

El registro de los datos fue llevado a cabo por los participantes en tablas tal como se muestran los datos de Adriana en la Figura 8.

Registro de resultados:

En la siguiente tabla registra tus resultados obtenidos al lanzar realmente los dados y sumar las caras superiores.

4	*2*	*6*	*7*	*4*	*3*	*12*	*10*	*9*	*10*	*7*	*10*	*7*	*4*	*5*	*6*	*6*	*5*	*4*	*6*

En las siguiente tabla registra la suma de las caras de los dados tus resultados inventados.

3	*12*	*2*	*11*	*3*	*10*	*4*	*9*	*11*	*6*	*7*	*6*	*7*	*5*	*7*	*5*	*4*	*8*	*8*	*4*

En la última tabla registra tus resultados obtenidos en la simulación del lanzamiento de los dados y suma de sus caras superiores, ordenados según su aparición en la tabal de frecuencias.

4	*4*	*5*	*5*	*5*	*6*	*6*	*6*	*6*	*6*	*7*	*8*	*8*	*8*	*9*	*9*	*9*	*10*	*10*	*12*

Figura 8. Evidencia de los registros de Adriana

Una vez que obtenían la información, se les pidió que contestaran una serie de preguntas tales como: ¿consideras que se puede distinguir una secuencia realmente aleatoria de otra que hemos inventado?, ¿tus resultados inventados se parecen a los obtenidos en el experimento o en la simulación?, y ¿qué sucede si en la aplicación el experimento se simula 1000 veces?

A la pregunta, ¿consideras que se pueden distinguir una secuencia realmente aleatoria de otra que hemos inventado?

Respuesta de los alumnos	**Alumnos**
No se puede distinguir las secuencias inventadas	Kenya, Nisa Danahe, Laura Selene, Iván, Adrián, Alfonso y Aniloreny
Si se puede distinguir las secuencias inventadas	Daniela, Victoria, Adriana, Brenda, Fernanda y Ángel

Figura 9. Respuesta de los alumnos a la pregunta, ¿consideras que se puede distinguir una secuencia realmente aleatoria de otra que hemos inventado?

Las respuestas de los 13 participantes se dirigen en dos vertientes, 7 de ellos consideran que definitivamente no hay forma de distinguir una secuencia aleatoria de otra donde los datos fueron inventados y los otros 6 piensan que si se puede distinguir una secuencia que se inventó como lo fue el caso de la invención de la información de los dados y otra aleatoria, pero reportan desconocer cómo hacerlo (Ver Figura 9)

Los alumnos al leer si ¿los resultados obtenidos son los mismos al realizar el lanzamiento de los dados, al inventar los datos, o al simular el experimento con una aplicación?

Todos coinciden que la información obtenida en los 3 sucesos, es diferente, como lo muestran las respuestas de Laura Selene y Ángel (Ver Figura 10).

Laura Selene: *No, los resultados no son los mismos.*

Ángel: A mí me salieron diferentes

Figura 10. Evidencias de Laura Selene y Ángel

El investigador realizó una intervención preguntando a los participantes, cuál había sido la forma o el camino que habían seguido al inventar los resultados de los dados. Después les pidió que le dictaran los datos de esa tabla, para poder hacer un análisis de los valores obtenidos (Anexo 2). La respuesta de los jóvenes es generalizada y refieren que escribieron números del 2 al 12, pero no pensaron como la suma de los dos dados. A continuación, se muestran los resultados de los alumnos:

Aniloreny:

7	5	2	11	10	3	5	12	9	6	7	4	5	12	8	6	4	5	6	8

Fernanda:

5	3	10	2	9	4	7	8	5	12	3	11	6	10	11	7	2	4	8	7

Figura 11. Evidencias de Aniloreny y Fernanda

Aniloreny y Fernanda externan que procuraron escribir todos los números al menos una vez, y demuestran un pensamiento que obedece a la creencia de la de la equiprobabilidad (Ver Figura 11).

Al cuestionarle a Nisa Danahe por qué había registrado 4 veces el número 7, comenta que simplemente se le ocurrieron. Lo mismo se les preguntó a Adrián, Adriana, Victoria, Brenda y Fernanda, que repitieron 3 veces el número 7 y la respuesta fue la misma, excepto Adrián que argumenta que el escribió varias veces el 7 porque es su número de la suerte.

Nisa Danahe:

5	*2*	*5*	*7*	*10*	*12*	*9*	*7*	*6*	*8*	*4*	*3*	*5*	*11*	*7*	*4*	*5*	*7*	*10*	*11*

Victoria:

3	5	7	10	2	3	5	5	8	11	12	10	9	8	7	10	4	6	7	11

Figura 12. Evidencias de Nisa Danahe y Victoria

En los registros de los datos inventados de estos alumnos se observa que realizaron la anotación de los datos obedeciendo al juicio intuitivo conocido como la heurística de la representatividad, ya que después de inventar por ejemplo un número, no repiten el número de manera inmediata, si no que tratan de espaciarlo. Además de que también estos participantes también asocian al experimento una equiprobabilidad (Ver Figura 12).

Al interrogar a Daniela, por qué no escribió el número 8, tampoco da una razón, ella explica que no se le ocurrió, además de que no da ninguna justificación del por qué repite el 9 de manera seguida (Ver Figura 13). Nuevamente se observa en el registro de Daniela, que sigue el juicio intuitivo de la representatividad y el sesgo de la equiprobabilidad.

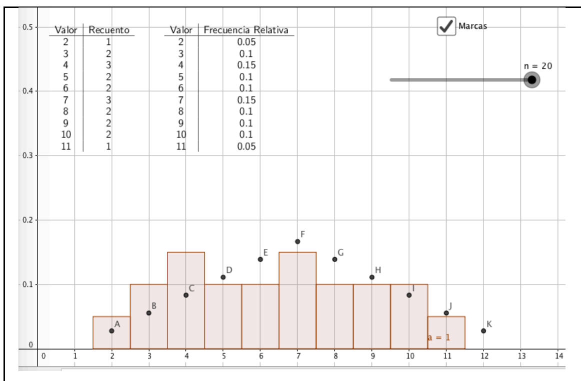
Daniela:

4	3	2	10	12	11	10	7	6	5	4	6	2	5	9	9	10	4	3	12

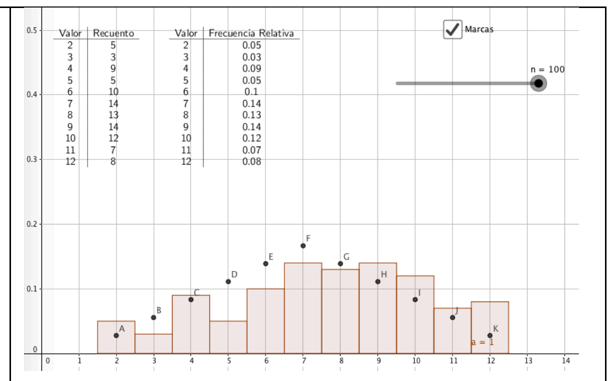
Figura 13. Evidencias de Nisa Danahe y Victoria

Es importante mencionar que ninguno de los alumnos reporta valores fuera del rango de 2 a 12. Al terminar la discusión el investigador les sugirió que tomaran nuevamente los dados de colores para que encontraran las opciones que existen de obtener el 5 como resultado de la suma. Los alumnos reflexionan y contestan los posibles pares (1, 4), (2, 3) (3, 2) y (4, 1).

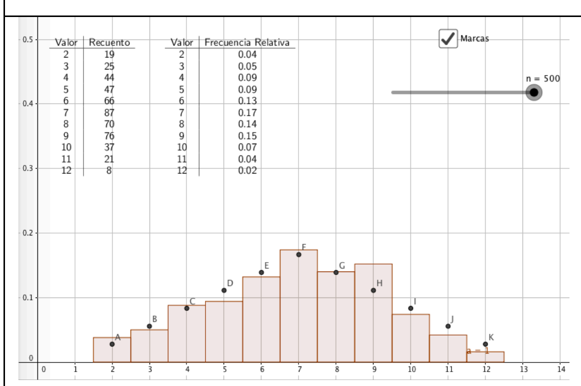
Antes de contestar la pregunta ¿ahora que sucede si se simula el experimento 1000 veces?, el investigador con ayuda del proyector les muestra cómo modificar en la aplicación el número de veces que se simula el lanzamiento de los dos dados, (Figura 14).



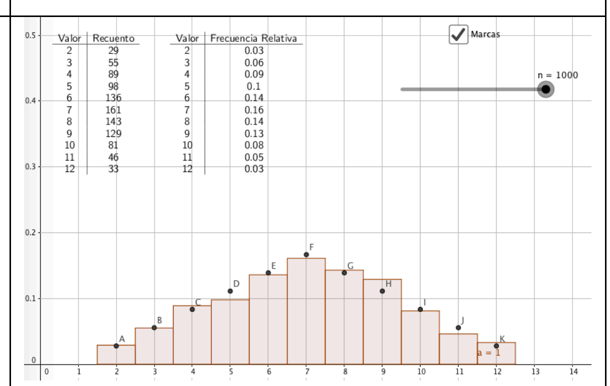
Simulación del evento 20 veces



Simulación del evento 100 veces



Simulación del evento 500 veces



Simulación del evento 1000 veces

Figura 14. Resultados de las simulaciones cuando “n” es 20, 100, 500 y 1000

El investigador les sugiere a los participantes que observen el comportamiento de la altura de las barras y los valores obtenidos, tanto en la tabla de frecuencias absolutas como en la tabla de frecuencias relativas. Las respuestas que escriben los estudiantes giran en torno a que entre mayor es el número del simulador, los valores de las tablas de frecuencias varían poco, que hay poco movimiento en la gráfica y reportan que se comporta de forma simétrica, como es el caso de Kenya y Fernanda. Victoria observa que cuando n = 1000, el valor más frecuente es 7 y es más estable, (Figura 15).

Kenya:

Las tablas de las frecuencias casi no cambian y hay poco movimiento en la gráfica

Fernanda:

Es casi siempre simetrica y no hay variación

Victoria:

La gráfica es más estable y el valor más frecuente es casi siempre el 7

Figura 15. Evidencias de Kenya, Fernanda y Victoria

El comportamiento que muestran las gráficas es el resultado de repetir el evento una gran cantidad de veces, estas siguen un patrón, el cual los alumnos observan y esto permite que se haga la conjetura de que entre más grande sea la simulación del experimento, el comportamiento que siguen las probabilidades será muy poco cambiante, generando una regularidad.

Una vez que los alumnos encuentran dicha regularidad en el comportamiento de las sumas de los posibles valores de las sumas de los dados al simular el experimento 1000 veces, el profesor solicitó a cada uno de los participantes los datos de los 3 eventos para obtener las Tablas 1, 2 y 3 y determinar el valor de la frecuencia de cada valor de la suma de las caras de los dados a la cual se llama frecuencia absoluta.

Tabla 1. Frecuencias absolutas de los datos reales

Número de suma al lanzar realmente los 2 dados	
Suma de puntos	Frecuencia absoluta
2	23
3	14
4	24
5	24
6	32
7	25
8	35
9	20
10	22
11	17
12	24
Total	260

Tabla 2. Frecuencias absolutas de los datos inventados

Número de la suma al inventar los valores de los 2 dados	
Suma de caras	Frecuencia absoluta
2	20
3	25
4	31
5	32
6	21
7	30
8	20
9	21
10	19
11	22
12	19
Total	260

Tabla 3. Frecuencias absolutas de los datos simulados

Número de suma al simular el lanzamiento de los 2 dados	
Suma de caras	Frecuencia absoluta
2	4
3	11
4	18
5	31
6	36
7	50
8	38
9	26
10	23
11	14
12	9
Total	260

El investigador solicitó a los estudiantes que las copiaran los datos de las Tablas 1, 2 y 3 en Excel y graficar cada una de ellas. Los estudiantes después de seleccionar cada una de las tablas eligieron de manera libre alguna de las opciones que aparecen en dicho software (columna agrupada, barra agrupada, dispersión con líneas rectas y línea); y responder: ¿el gráfico que escogiste te permite comprender y resumir los datos de la suma de las caras superiores obtenidos en los tres casos (experimento, datos inventados y simulación)?, ¿cuál es el valor más frecuente en cada uno de los casos?, ¿en qué se parecen?, ¿en qué se diferencian?, ¿el valor más frecuente es el mismo?, ¿piensas que la forma que inventaste los valores de los dados que se obtendrían al lanzar 20 veces un par de dados legales es totalmente acertada?

En la pregunta, ¿La gráfica que escogiste te permite comprender y resumir los datos de la suma de las caras superiores obtenidos en los tres casos (experimento, datos inventados y simulación)?

Los alumnos tuvieron la oportunidad de ver los diferentes tipos de gráficas que ofrece Excel; una vez que vieron las diferentes opciones, ellos tomaron la decisión de elegir la que a su criterio consideraban les ayudaba a comprender y representar los datos obtenidos en cada uno de los tres casos. Algunos estudiantes dejaron el aspecto de las gráficas que proporciona de manera predefinida Excel, y otros decidieron cambiar la apariencia inicial, modificando el color y agregando etiquetas, al hacer estos cambios les pareció más atractiva y útil la información. Por ejemplo, Iván menciona que el eligió la gráfica de columnas verticales, porque dentro de cada barra aparece el número de veces que suceden las sumas de los dados. Kenya, Laura Selene, Iván, Brenda, Alfonso, Aniloreny y Ángel eligieron la gráfica de columnas verticales, Nisa Danahe, Victoria y Fernanda escogieron la gráfica de columnas horizontales, mientras que Daniela, Adrián y Adriana optaron por la gráfica de líneas, ver Figuras 16, 17 y 18, respectivamente.

Las respuestas que reportan todos los alumnos cuando se les cuestiona, ¿cuál es valor más frecuente en cada uno de los casos? Logran identifican correctamente los valores más frecuentes de cada uno de los casos. Siendo estos 35 para los lanzados, 32 para los inventados y 50 para los simulados. Incluso algunos de ellos mencionaron el valor más frecuente y el menos frecuente, como Victoria y Daniela (Ver Figura 19)

Victoria: Para la primera 14 y 35, para la segunda 19 y 32 y en la última 4 y 50

Daniela: En los lanzados es 35, para los inventados 32 y para los simulados 50.

Figura 19. Evidencias Victoria y Daniela

Posteriormente se les pide que comparen los tres gráficos y respondan a las preguntas, ¿en qué se parecen?, ¿en qué se diferencian?, ¿el valor frecuente es el mismo?

Los alumnos reportan que los tres gráficos obtenidos no se parecen, reiterando en la siguiente pregunta que las tres gráficas son totalmente diferentes.

Además, nuevamente todos los participantes coinciden en señalar que los valores más frecuentes de cada caso son diferentes, algunos participantes reportan el comportamiento de uno de las gráficas, por ejemplo, Laura Selene menciona que la gráfica de los datos simulados tiene apariencia de triángulo, mientras que Daniela escribe que en las dos primeras existen varios aumentos y disminuciones de frecuencias, mientras que en la gráfica de los simulados solo una vez hay un aumento y disminución (Ver Figura 20).

Laura Selene: *Entre ellas no se parecen, pero la última gráfica parace un triángulo.*

Daniela: Las dos primeras gráficas existen subidas y bajadas, pero en la última desde 2 hasta la mitad sube y después solo baja.

Figura 20. Evidencias de Laura Selene y Daniela

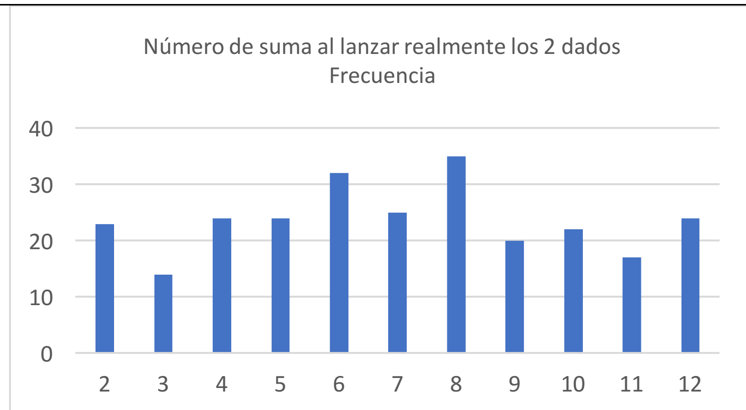
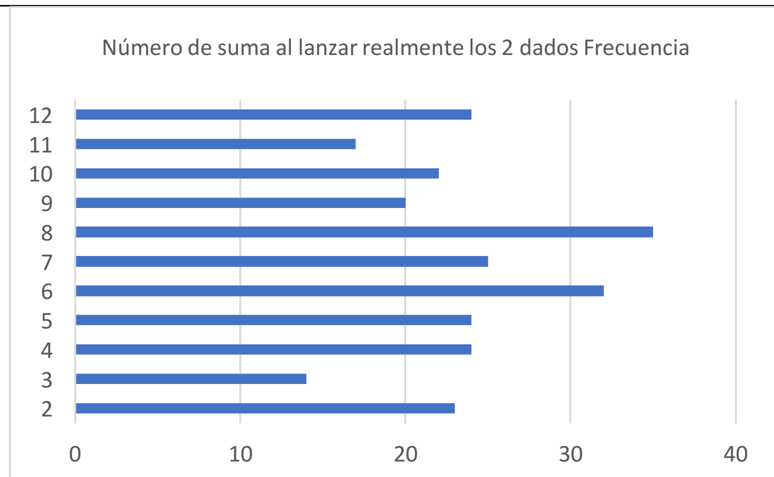
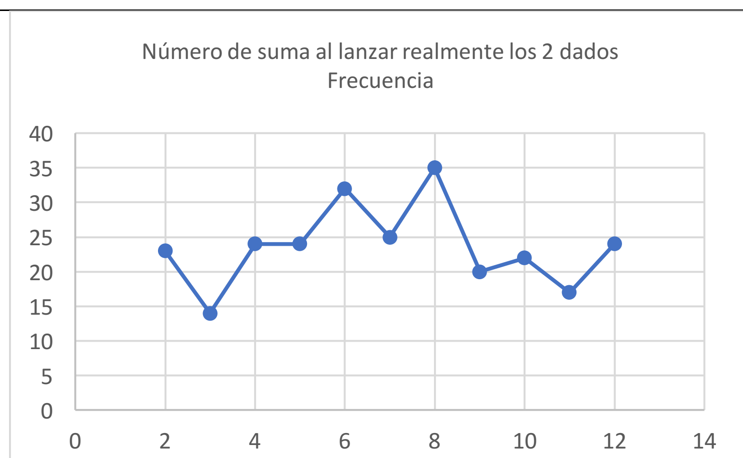
Alumnos	**Tipo de gráfico**	**Gráfico**
Kenya, Laura Selene, Iván, Brenda, Alfonso, Aniloreny y Ángel	Columnas verticales	
Nisa Danahe, Victoria y Fernanda	Columnas horizontales	
Daniela, Adrián y Adriana	Líneas	

Figura 16. Evidencias de gráficos construidos por los alumnos para los dados lanzados

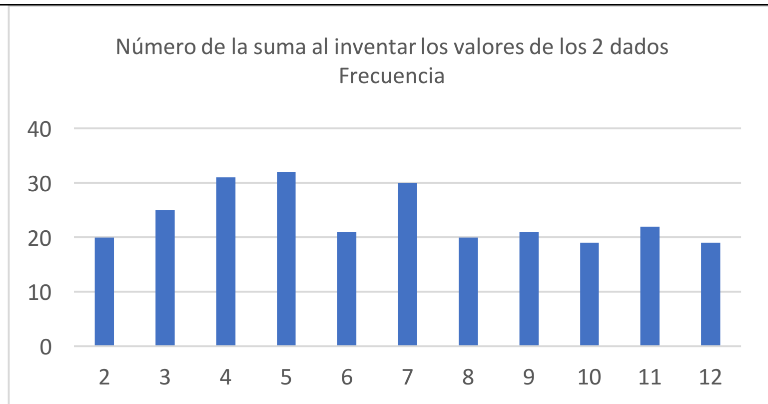
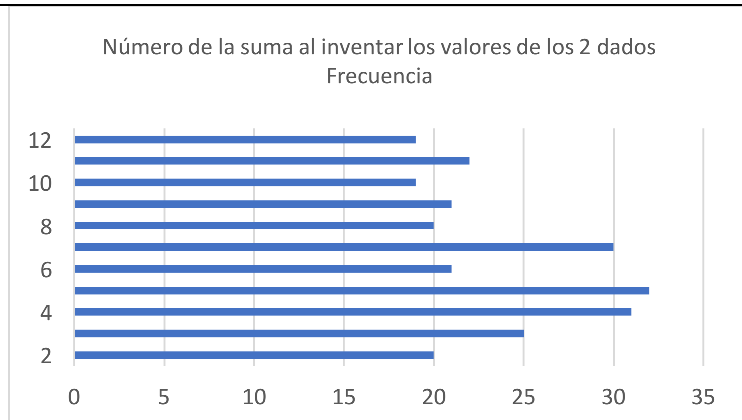
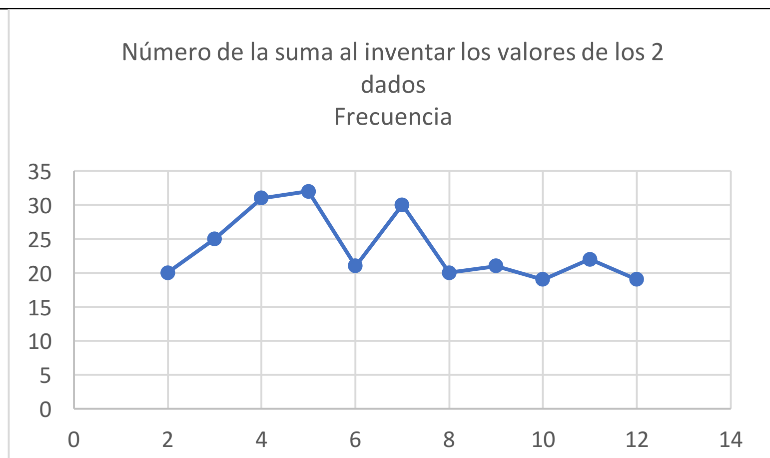
Alumnos	Tipo de gráfico	Gráfico
Kenya, Laura Selene, Iván, Brenda, Alfonso, Aniloreny y Ángel	Columnas verticales	
Nisa Danahe, Victoria y Fernanda	Columnas horizontales	
Daniela, Adrián y Adriana	Líneas	

Figura 17. Evidencias de gráficos construidos por los alumnos para los dados inventados

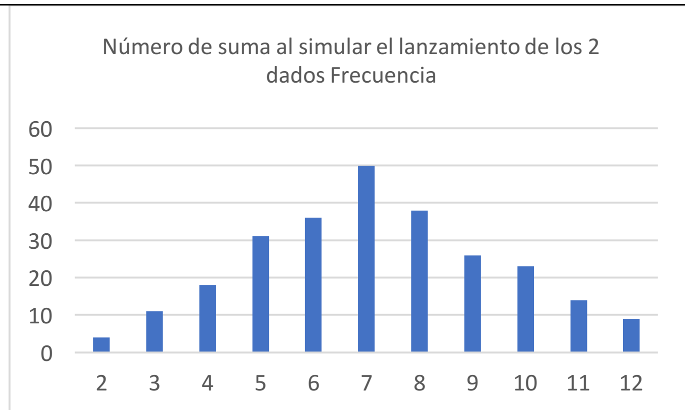
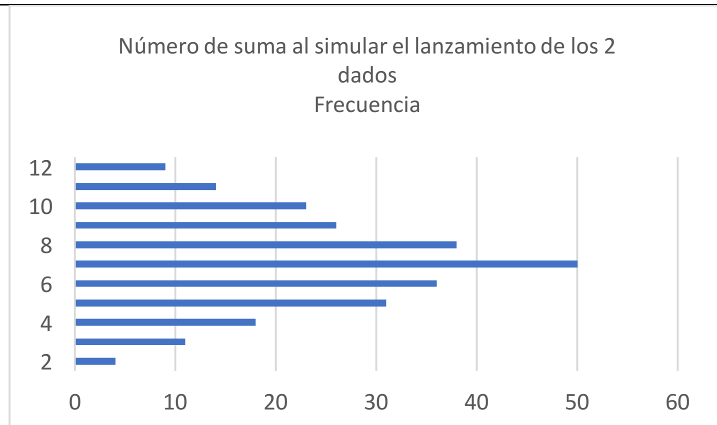
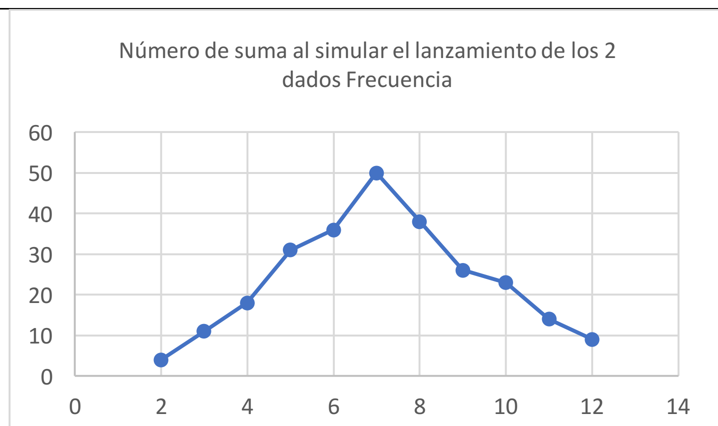
Alumnos	Tipo de gráfico	Gráfico
Kenya, Laura Selene, Iván, Brenda, Alfonso, Aniloreny y Ángel	Columnas verticales	
Nisa Danahe, Victoria y Fernanda	Columnas horizontales	
Daniela, Adrián y Adriana	Líneas	

Figura 18. Evidencias de gráficos construidos por los alumnos para los dados simulados

Cuando los participantes concluyeron la elaboración de las gráficas, el investigador cuestionó a los estudiantes por qué eligieron ese tipo de gráficas, Laura Selene mencionó que al utilizar gráficas de barras verticales se hace evidente de manera visual el valor de mayor frecuencia en el centro (vértice del triángulo que menciona ella), (Ver Figura 21).

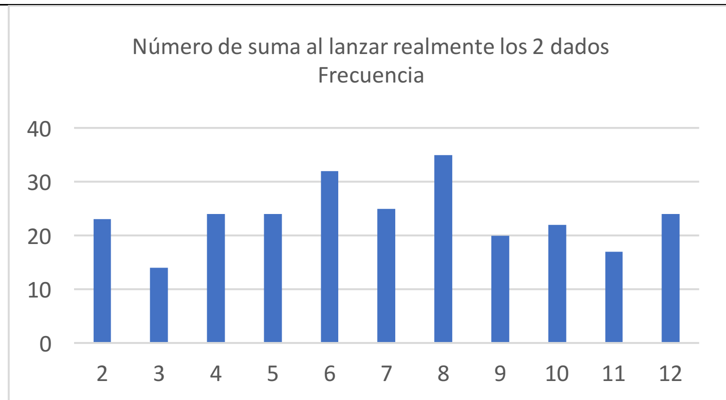
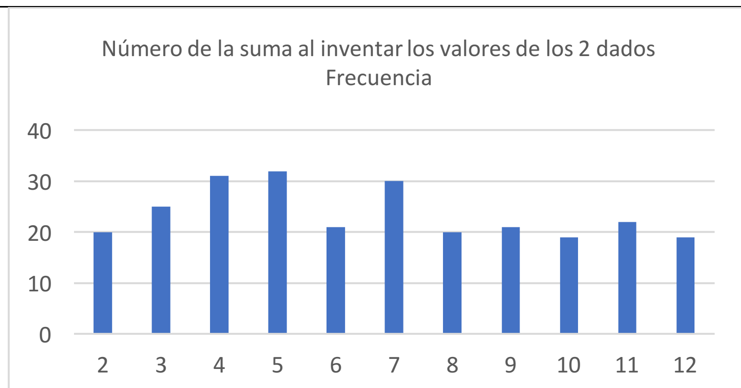
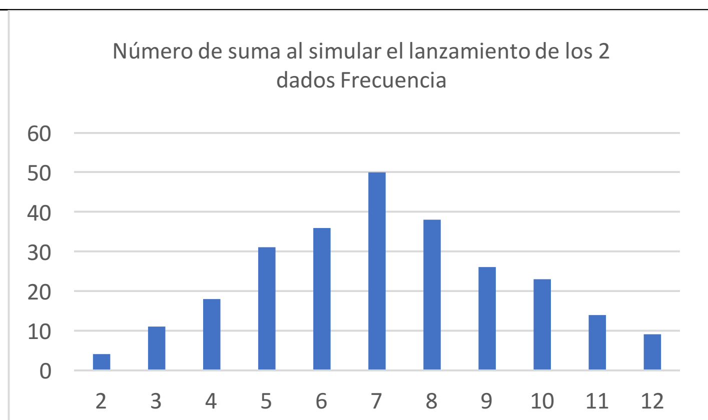
Gráfica de los dados lanzados	
Gráfica de los dados inventados	
Gráfica de los dados simulados	

Figura 21. Comparativo de las gráficas de Laura Selene

Después de observar las gráficas, los alumnos contestaron la pregunta, ¿consideras que la forma cómo pensaste los números que se obtienen al inventar 20 veces un par de dados legales es totalmente correcta?

La intención de esta pregunta era hacer reflexionar a los participantes con respecto a lo que su forma de inventar los valores de los dados, haciendo una comparación entre los datos inventados y los simulados. La mayoría de los estudiantes reportan que a la hora de inventar los valores de los datos no pensaron en los pares que forman los dados, solo pensaron en la suma de las caras y que solo inventaron los números conforme se les ocurría. Este pensamiento evidencia que los participantes valoraron las probabilidades de cada valor de la suma de forma equiprobable.

Daniela y Kenya realizan señalamientos que en la gráfica de los inventados su apariencia plana, debido a que los valores de las frecuencias de los valores de los extremos (2 y 12) son similares a las de los valores centrales (6, 7 y 8) (ver figura 22). Evidenciando de esta forma el sesgo de equiprobabilidad, ahora visto de una forma gráfica (ver Figura 23).

Daniela:

> Veo que las frecuencias cambian mucho en los dados inventados, y en los simulados los valores del centro son mayores a los de los extremos. Entonces creo que mi intuición no es tan correcta.

Figura 22. Evidencia de Daniela

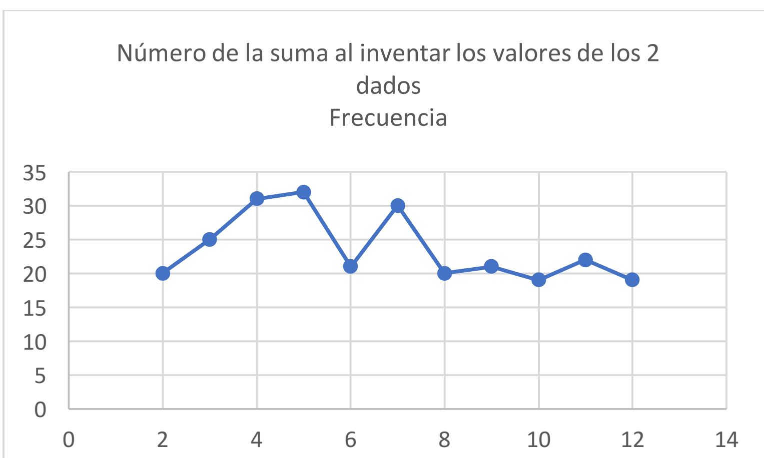


Figura 23. Gráfica de los dados inventados obtenida por Daniela

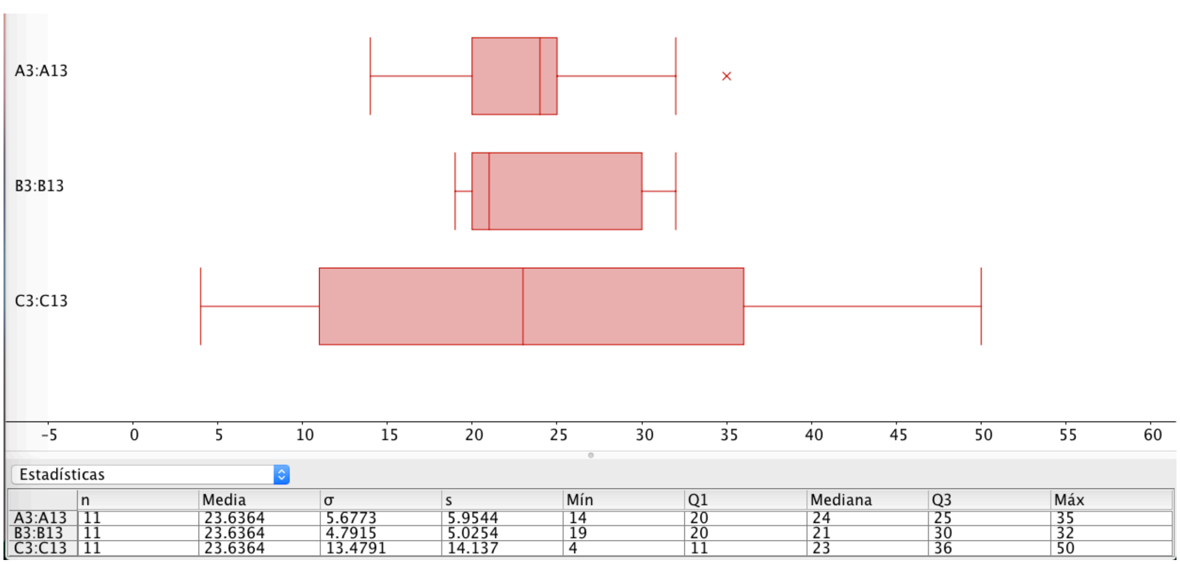
Laura Selene y Brenda, indican que su forma de pensar si es acertada, pero no escriben ninguna justificación.

Se construyó una tabla con todos los datos obtenidos al lanzar, inventar y simular el lanzamiento de los dados de manera simultánea (Figura 24), para después se determinar las medidas de tendencia central y dispersión.

Tabla 4. Frecuencias absolutas de los 3 eventos

Suma de caras superiores	**Frecuencia real**	**Frecuencia inventada**	**Frecuencia simulada**
2	23	20	4
3	14	25	11
4	24	31	18
5	24	32	31
6	32	21	36
7	25	30	50
8	35	20	38
9	20	21	26
10	22	19	23
11	17	22	14
12	24	19	9

Posteriormente, los participantes graficaron en GeoGebra un diagrama de caja y bigotes, con los resultados de la Tabla 4 (Ver Figura 24) y así contestar las preguntas: ¿tienen alguna medida de tendencia central (media, mediana, moda), igual?



	n	Media	σ	s	Mín	Q1	Mediana	Q3	Máx
A3:A13	11	23.6364	5.6773	5.9544	14	20	24	25	35
B3:B13	11	23.6364	4.7915	5.0254	19	20	21	30	32
C3:C13	11	23.6364	13.4791	14.137	4	11	23	36	50

Figura 24. Gráfica del diagrama de caja y bigotes con los datos de la Tabla 4

Todos los participantes al observar la gráfica del diagrama de caja y bigotes y registran de manera satisfactoria que los valores de las medianas en los tres casos son diferentes. Kenya, Nisa Danahe, Daniela, Laura Selene, Iván, Victoria y Adriana además reportan los valores 24, 21 y 23, de las medianas respectivamente. También visualizan que el único valor de tendencia central que es igual en los tres sucesos es la media cuyo valor es igual a 23.6364 (Figura 25)

Kenya: Solo la media es la misma, 23.6364, las modas son diferentes y son iguales a los valores más frecuentes

Laura Selene: *La media para las tres es 23.6364, pero las medianas si son diferente en cada caso, los valores son 24, 21 y 23.*

Victoria: La mediana en todas es diferente y sus valores son 24, 21 y 23, pero la media igual en todas

Figura 25. Evidencias de Kenya, Laura Selene y Victoria

En la discusión de los resultados, el investigador pregunto qué significa la X, que aparece en la gráfica de los datos los dados lanzados. Daniela y Kenya, comentan que era un valor de la tabla y surgió la pregunta, ¿cuál es ese valor X? Para identificar el valor, el investigador les sugirió que observaran el eje de las abscisas y buscaran, ¿cuál era el valor de la suma de las caras de los dados lanzados que tenía una frecuencia aproximadamente de 35?, la respuesta de los jóvenes es 8. Para continuar la discusión se les solicitó que cambiaran ese valor en GeoGebra, aumentándolo y disminuyéndolo, para ver qué sucedía en la gráfica. Los alumnos identifican que si el valor se aumenta la X se mueve a la derecha, mientras que si lo disminuyen el valor se mueve a la izquierda y si se coloca un número menor o igual a 33, la X desaparece. Al terminar de hacer conjeturas los alumnos, el investigador les comenta esa X, es un *valor atípico* o *valor aislado* y en ocasiones se presentan debido a errores de medición o registro, como pudo haber sido el caso de nuestros datos, ya que los dados son comunes (no calibrados) y la forma y el lugar donde realizaron el lanzamiento de los dados, debido a que algunos los lanzaron al suelo y otros al escritorio. Y aunque este no era un tópico considerado en la secuencia, la herramienta permitió que los alumnos conocieran el concepto de valor atípico.

Para concluir este episodio se dio paso a la última parte de la actividad, en la cual los participantes debían mover varias veces el deslizador que modificaba el número de simulaciones hasta n = 1000 en el modelo dinámico construido en GeoGebra, para que nuevamente observaran su comportamiento y después se dieran a la tarea de investigar en internet los conceptos de distribución de probabilidad y esperanza matemática, y finalmente contestaran preguntas como: si el número de simulaciones en la aplicación en Geogebra fuese por ejemplo n = 1000, ¿cuántas veces crees que se vayan a obtener el número 7?, si lanzas un par de dados distinguibles (diferente color) una sola vez, ¿cuántas veces puede aparecer 7 y cuáles serían los pares ordenados?, en términos de probabilidad, ¿cuál será la probabilidad para que puede aparecer un 7 al realizar un solo lanzamiento el par de dados?, y ¿si simulas 1000 veces el lanzamiento del par de dados?, ¿cuál es la distribución de probabilidad del lanzamiento de un par de dados legales?, calcula ésta para el lanzamiento de un par de dados.

Si el número de simulaciones en la aplicación en GeoGebra fuese por ejemplo n = 1000, ¿cuántas veces crees que se vayan a obtener el número 7?

Las respuestas que reportan todos los alumnos son diferentes, por ejemplo, Nisa Danahe 170, Daniela 169 y Victoria 172 (Figura 26)

Nisa Danahe: *Al mover el número del deslizador obtuve 170*

Daniela: Si muevo el deslizador el resultado es 169.

Victoria: Cuando cambie el valor de n, me salió 172

Figura 26. Evidencias de Nisa Danahe, Daniela y Victoria

El investigador les solicita de manera verbal que digan los valores que obtuvieron en voz alta y las respuestas son todas diferentes, surgiendo la pregunta ¿por qué sucede esto? Los alumnos argumentan a que la aplicación simula el lanzamiento de los dados y que si se arroja un par de dados las respuestas los resultados de las caras son diferentes en cada lanzamiento. Esta es una conclusión importante ya que los participantes observan que la aplicación en GeoGebra simula el evento, el cual se realiza en condiciones similares, pero no iguales, ya que, si esto sucediera, entonces no sería un evento aleatorio.

Si lanzas un par de dados distinguibles (diferente color) una sola vez, ¿cuántas veces puede aparecer 7 y cuáles serían los pares ordenados?

Antes de contestar la pregunta, el investigador les sugiere que nuevamente tomen los dados con una mano cada uno y los muevan ambos de forma ordenada. De esta forma determinan las 36 parejas de números al lanzar ambos dados (espacio muestral). Esta pregunta buscaba que los alumnos fueran capaces de identificar y reportar los seis puntos muestrales cuya suma es igual a 7. Todos los participantes reportan correctamente el resultado a esta pregunta; los alumnos escriben las seis diferentes formas en que puede aparecer el 7, que son (1, 6), (2, 5), (3, 4), (4, 3), (5, 2) y (6, 1)

Después contestaron, ¿cuál será la probabilidad para que puede aparecer un 7 al realizar un solo lanzamiento el par de dados?, y ¿si simulas 1000 veces el lanzamiento del par de dados?

Antes de contestar la pregunta les sugiere que una vez que tengan el valor de la probabilidad y el valor obtenido en la aplicación de GeoGebra, les sugiere que realicen la comparación de ambas. Los participantes indican que sus resultados son diferentes pero muy cercanos, como lo escribe Kenya, Iván y Victoria (Figura 27)

Kenya:
Probabilidad p = 6/36 = 0.1666
Simulación p = 164/1000 = 0.164
Los resultados son muy parecidos

Iván: *La probabilidad es 6/36, y con el programa el resultado es 175/1000. De hecho, si hago las divisiones son casi iguales*

Victoria: En el simulado es 172/1000 y con los pares 6/36

Figura 27. Evidencias de Kenya, Iván y Victoria

La intención de esta pregunta era que los participantes compararan el valor de la probabilidad de obtener un 7 al lanzar un par de dados y sumar las caras superiores y lo compararan con el valor que se obtiene en la aplicación hecha en GeoGebra cuando n = 1000.

A continuación, se les solicito a los alumnos que investigaran el concepto de distribución de probabilidad para contestar la pregunta, ¿cuál es la distribución de probabilidad del lanzamiento de un par de dados legales?

El investigador les reitera a los participantes que es importante que incluyan la dirección electrónica de la página de consulta, se muestran algunas definiciones encontradas por Kenya, Daniela, Nisa Danahe y Adriana (Ver Figura 28)

Kenya:

En teoría de la probabilidad y estadística, la **distribución de probabilidad** de una variable aleatoria es una función que asigna a cada suceso definido sobre la variable aleatoria la probabilidad de que dicho suceso ocurra. La distribución de probabilidad está definida sobre el conjunto de todos los sucesos y cada uno de los sucesos es el rango de valores de la variable aleatoria. También se dice que tiene una relación estrecha con las distribuciones de frecuencia. De hecho, se puede entender que una distribución de probabilidades sería una frecuencia teórica, ya que ésta última es aquella que describe cómo se espera que varíen los resultados.

https://es.wikipedia.org/wiki/Distribuci%C3%B3n_de_probabilidad

Daniela:

Una distribución de probabilidad la podemos concebir como una distribución teórica de frecuencia, es decir, es una distribución que describe como se espera que varíen los resultados. Dado que esta clase de distribuciones se ocupan de las expectativas son modelos de gran utilidad para hacer inferencias y tomar decisiones en condiciones de incertidumbre.

Nisa Danahe:

*Sea X una variable aleatoria discreta cuyos valores suponemos ordenados de menor a mayor. Llamaremos **función de distribución de la variable X**, y escribiremos F(x) a la función:*

$$F(x) = p(X \leq x)$$

La función de distribución asocia a cada valor de la variable aleatoria la probabilidad acumulada hasta ese valor.

http://www.vitutor.com/pro/3/a_3.html

Adriana:

Las distribuciones de probabilidad son de tipo discreto y continuo, según la variable aleatoria que este en cuestión, luego en este aparte se estudiaran dichas distribuciones, sus principios, la función que la identifica, sus propiedades y los campos de aplicación de las mismas.

http://www.wikillerato.org/Esperanza_matem%C3%A1tica.html

Figura 28. Evidencias de Kenya, Daniela, Nisa Danahe y Adriana

Una vez que los alumnos investigaron el concepto solicitado, procedieron a calcular la distribución de probabilidad del lanzamiento y suma de las caras superiores de un par de dados.

Kenya, Laura Selene, Victoria, Adriana, Fernanda, Iván y Aniloreny, reportan con éxito la distribución de probabilidades (Ver Figura 29). Estos alumnos empleando la definición y el espacio muestral de los 36 pares que surgen del lanzamiento de los dados, son capaces de encontrar las funciones de distribución de las probabilidades. En otras palabras, estos participantes son capaces de relacionar el valor de probabilidad asignado a cada posible valor de la suma de las caras, lo cual contribuye a modificar el pensamiento que tenían en un inicio de la actividad, en donde mencionaban que un evento aleatorio no obedecía a ninguna ley o patrón establecido.

Kenya:

2	3	4	5	6	7	8	9	10	11	12
1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Fernanda:

2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

Adriana:

Suma	2	3	4	5	6	7	8	9	10	11	12
Valor	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Figura 29. Evidencias de Kenya, Fernanda y Adriana

Daniela escribe los valores de la distribución de frecuencias relativas que se despliega en la aplicación de GeoGebra (Ver Figura 30). Esto aunque es un error, ya que no es la función de distribución de probabilidad, permite observar que la alumna relaciona la probabilidad de un evento, con la frecuencia de su ocurrencia, al simular muchas veces (en este caso 1000 veces) un experimento aleatorio.

Daniela:

Estos valores son mis valores:

Valor	Frecuencia relativa
2	0.02
3	0.05
4	0.1
5	0.1
6	0.12
7	0.17
8	0.14
9	0.11
10	0.1
11	0.05
12	0.03

Figura 30. Evidencia de Daniela

Nisa Danahe, Adrián, Brenda, Alfonso y Ángel, no reportan resultados, al ser cuestionados mencionan que no comprendieron la definición encontrada en internet, debido a que existen palabras y frases que no conocían, como: variable aleatoria discreta, variable aleatoria continua, función de probabilidad acumulada, valor de asignación. El investigador aprovecha para invitar a los alumnos a que sean críticos con la información que encuentran en internet y no se queden con lo primero que se despliega en el buscador.

Por último, se les solicitó a los alumnos que observaran y compararan los valores de las frecuencias relativas de la aplicación en GeoGebra cuando el deslizador es n = 1000 y la distribución de probabilidad encontrada en el inciso anterior; y contestaran la pregunta, si se parecen o difieren, ¿qué puedes decir al respecto?

Los siete participantes que reportaron la distribución de probabilidad correctamente son capaces de contestar esta pregunta, y argumentan ambas distribuciones son muy parecidas o casi iguales, como lo escriben Kenya, Laura Selene y Fernanda (Ver Figura 31)

> Kenya: Si hago la división de los quebrados, los valores son casi iguales a los obtenidos con el GeoGebra
>
> Laura Selene: *Los valores del GeoGebra son casi iguales a los valores de las probabilidades, solo que las probabilidades están en fracciones*
>
> Fernanda: *Después de dividir los quebrados y compararlos con el GeoGebra, veo que, si son casi iguales, la diferencia son décimas*

Figura 31. Evidencias de Kenya, Laura Selene y Fernanda

De esta forma y con la ayuda de la simulación del experimento construida en GeoGebra, se encontró una conexión entre la distribución de frecuencias (cuando el experimento se simula muchas veces y la distribución de probabilidad), es decir la relación entre un concepto intuitivo y un modelo teórico. Además de que el grupo de participantes en conjunto con el investigador se logró una reflexión de que al enfrentarse a un evento aleatorio definitivamente no es posible saber el resultado del mismo de manera precisa, ya que éste está sujeto al azar, pero si sigue un comportamiento, el cual puede orientarnos a saber cuáles son los valores que tienen una mayor y una menor probabilidad de ocurrencia, si el experimento se realiza un número considerable de veces.

4.2.1 Análisis de la segunda actividad

Al cuestionarle a los alumnos por la forma que siguieron al determinar los valores de los dados inventados, se observa que siguieron lo que Kahneman y Tversky (1973), llaman la heuristica de la representatividad, debido a que solo obedecieron a su intuitición al pensar en números y dejaron de lado, de dónde se obtienen esos números. Los participantes mencionan que escribieron al menos una sola vez los números de 2 al 12, lo cual es complicado que suceda en 20 lanzamientos, estimando de esta manera la probabilidad de que todos los valores debían tener la misma probabilidad de salir. No reflexionaron en el hecho que esos números, provienen de la suma de los valores de un par de dados y hubo necesidad de intervenir en la aplicación de la actividad, para mostrarles que si se obtiene como resultado un 5, las opciones de las posibles sumas son: 1+4, 2+3, 3+2 y 4+1.

Las respuestas que reportan los alumnos cuando modifican con el deslizador de la aplicación en GeoGebra el número de veces que se simula el lanzamiento de los dados de n=20 a n=1000, muestran una evolución en la concepción azarosa de la probabilidad, argumentando que existe un comportamiento no definido muy bien definido cuando n = 20, y comienza a estabilizarse las frecuencias de las gráficas y los valores de las tablas cuando se aumenta el número de simulaciones (Figura 14), de esta forma se promueve el desarrollo de una idea intuitiva de la distribución de probabilidades del experimento. Es importante mencionar que el deslizador de la herramienta digital fue el potenciador de las ideas mostradas por los participantes, dando la facilidad de cambiar el tamaño de las simulaciones y de esta forma mostrar los comportamientos de las frecuencias al variar el número de las simulaciones. Y de esta forma los alumnos pudieron realizar la conjetura de que un experimento aleatorio, si se realiza muchas veces, este tiende a seguir un comportamiento en sus probabilidades para cada valor asignado en este caso a la suma de las caras de los dados.

Al realizar la representación gráfica de los datos, los alumnos enuncian criterios para seleccionar aquellas opciones que consideraron más pertinentes para dar una buena imagen de los datos que están manejando y consiguieron identificar algunos errores que suelen cometerse en la elaboración y selección de las gráficas (Figuras 16, 17 y 18). Se resalta la importancia de Excel y GeoGebra como herramientas para elaborar las gráficas de una manera sencilla, pero precisa. Fue relevante discutir con los estudiantes que algunos programas son una mejor opción al elegir el tipo de gráfica que se pretende hacer. Por ejemplo, GeoGebra es una mejor opción que Excel para realizar la gráfica de caja y bigotes (Figura 24), ya que cuenta con una opción dentro del menú de análisis de datos multivariados, en donde además se despliegan las medidas de tendencia central de los eventos estudiados. Así además se puede orientar a los alumnos a reconocer la utilidad de este tipo de medidas para poder hacer comparativos de los comportamientos de los datos de cada una de las 3 tablas. Una situación que no se tenía contemplada fue la aparición de un valor atípico en la gráfica de caja y bigotes de la distribución de los datos, pero la tecnología permitió, extender la discusión y permitir a los estudiantes conjeturar y encontrar cuál era ese valor atípico.

Los alumnos evidencian la existencia de una similitud entre el valor de la frecuencia relativa que se obtiene en la aplicación y la probabilidad al realizar la división de los casos exitosos (puntos muestrales) entre todos los casos posibles (espacio muestral) (Figura 27). En esta tarea es relevante el uso de la aplicación de GeoGebra, debido a que como se escribió con anterioridad, el deslizador les permite a los alumnos obtener un número considerable de simulaciones del evento (Figura 26).

El internet es una herramienta útil para buscar información de conceptos estadísticos como lo comentan Borba, Clarkson, y Gadanidis (2013), pero es pertinente mencionar que los alumnos deben de estar consientes y preparados para ser críticos con la información encontrada en la red, debido a que no toda la información que circula en internet es cierta o útil a los propósitos requeridos. Esta situación queda evidenciada al solicitarles a los alumnos que busquen la definición de distribución de probabilidad; ellos reportan la respuesta de manera exitosa, pero al pedirles que hagan uso de esa definición, no fueron capaces de codificar y comprender lo reportado en su definición, en consecuencia, no encuentran lo solicitado. Al revisar detenidamente las definiciones de los alumnos, como ellos le comentaron al investigador, existe información, términos y simbología matemática, que pudo ser determinante para que no tuvieran una comprensión de la definición.

Los alumnos que determinaron la distribución de probabilidad (Figura 29), son capaces de entender la definición encontrada en internet, asignándole el valor de la probabilidad a cada posible valor obtenido de la suma de las caras de los dados (valores de la variable aleatoria discreta). También son capaces de encontrar la similitud que existe entre la distribución de frecuencias relativas en la aplicación de GeoGebra que describe la forma cómo se espera que varíen los resultados y la distribución de probabilidad, que arroja los valores teóricos de las probabilidades de los eventos (Figura 31).

4.3 Tercera actividad – Ajuste de una línea recta

En la quinta sesión antes de iniciar la actividad, el investigador le pregunto a los alumnos, ¿si conocían alguna dependencia o institución que posea información (fidedigna) de la población de México?

El INEGI y el INE, fueron las dos únicas dependencias gubernamentales que los alumnos identificaron. Entonces se propuso que buscaran en internet, algunas otras opciones, para hacer un compilado de las entidades encontradas, con información de la población de México (Figura 32).

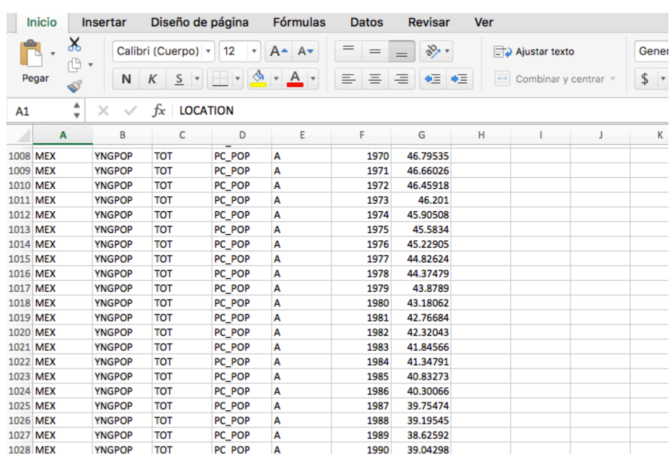
-SNIEG (Sistema Nacional de Información Estadística y Geográfica) que tiene el propósito de difundir información de interés Nacional. (www.snieg.mx)

*-El Consejo Nacional de Población (**CONAPO**), que es una instancia gubernamental mexicana que tiene por objeto regular el crecimiento de la población, los movimientos demográficos, así como la distribución de los habitantes de México en el territorio (http://www.conapo.gob.mx/ES/CONAPO/Mexico_en_cifras)*

-ENADID (Encuesta Nacional de la Dinámica Demográfica), es una institución que brinda estadísticas sobre fecundidad, mortalidad infantil y movimientos migratorios. (http://www.inegi.org.mx/est/contenidos/proyectos/encuestas/hogares/especiales/enadid/)

Figura 32. Instituciones que cuentan con información de la población de México

El investigador compartió con los estudiantes la información de la población joven de los países miembros de la OCDE con la que se iban a trabajar en la actividad. El archivo de Excel contiene: 1) el nombre del país, 2) el indicador (Young population – Población joven), 3) tamaño de la población considerada – Total, 4) la medida – Porción de población, 5) la frecuencia – Anual, 6) el año y 7) el valor del indicador (Figura 33). Los alumnos únicamente seleccionaron los datos de México.



	A	B	C	D	E	F	G
1008	MEX	YNGPOP	TOT	PC_POP	A	1970	46.79535
1009	MEX	YNGPOP	TOT	PC_POP	A	1971	46.66026
1010	MEX	YNGPOP	TOT	PC_POP	A	1972	46.45918
1011	MEX	YNGPOP	TOT	PC_POP	A	1973	46.201
1012	MEX	YNGPOP	TOT	PC_POP	A	1974	45.90508
1013	MEX	YNGPOP	TOT	PC_POP	A	1975	45.5834
1014	MEX	YNGPOP	TOT	PC_POP	A	1976	45.22905
1015	MEX	YNGPOP	TOT	PC_POP	A	1977	44.82624
1016	MEX	YNGPOP	TOT	PC_POP	A	1978	44.37479
1017	MEX	YNGPOP	TOT	PC_POP	A	1979	43.8789
1018	MEX	YNGPOP	TOT	PC_POP	A	1980	43.18062
1019	MEX	YNGPOP	TOT	PC_POP	A	1981	42.76684
1020	MEX	YNGPOP	TOT	PC_POP	A	1982	42.32043
1021	MEX	YNGPOP	TOT	PC_POP	A	1983	41.84566
1022	MEX	YNGPOP	TOT	PC_POP	A	1984	41.34791
1023	MEX	YNGPOP	TOT	PC_POP	A	1985	40.83273
1024	MEX	YNGPOP	TOT	PC_POP	A	1986	40.30066
1025	MEX	YNGPOP	TOT	PC_POP	A	1987	39.75474
1026	MEX	YNGPOP	TOT	PC_POP	A	1988	39.19545
1027	MEX	YNGPOP	TOT	PC_POP	A	1989	38.62592
1028	MEX	YNGPOP	TOT	PC_POP	A	1990	39.04298

Figura 33. Archivo de Excel con la información de la población joven

Cuando todos los alumnos tuvieron a su disposición los datos históricos de la población joven en México, iniciaron a una sesión en GeoGebra en donde se los solicitó que eligieran la opción de hoja de cálculo, para que pudieran exportar los datos del archivo de Excel, colocando el año en la columna A (el eje X) y el valor de la población joven en México en la columna B (el eje Y); cuando era seleccionada esta información en GeoGebra se despliega un menú, en donde los participantes debían escoger la opción de "Análisis de regresión de dos variables", una vez realizada esta acción la herramienta generaba un diagrama de dispersión o nube de puntos, el cual representa los datos de una distribución estadística bidimensional mediante coordenadas cartesianas (Figura 34).

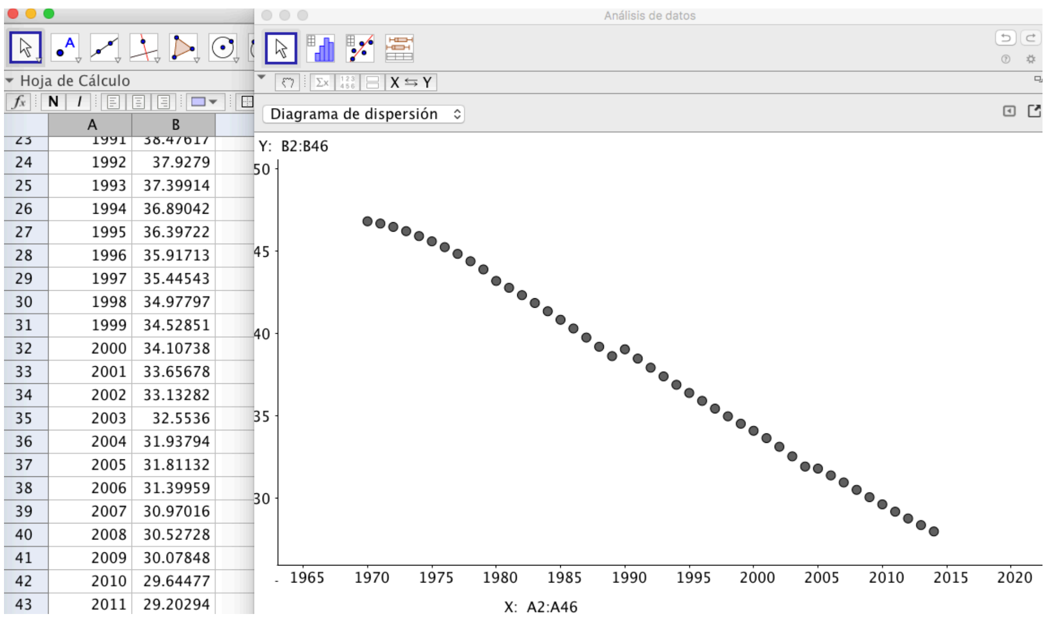


Figura 34. Diagrama de dispersión o nube de puntos

Una vez que los participantes obtuvieron la gráfica de dispersión de los datos en GeoGebra, el investigador preguntó, ¿qué forma geométrica aparentan seguir los datos graficados? La respuesta de los alumnos fue: una línea recta. Y aprovechando esa respuesta se les solicitó que construyeran un par de deslizadores, los cuales iban a controlar el movimiento de los parámetros pendiente m y la ordenada al origen de una recta b; para crear una familia de rectas y de esta forma realizaran un ajuste visual de la recta, registrando los valores que ellos consideraran más pertinentes.

Los estudiantes enfrentaron al problema de tener que identificar los intervalos en los que tenían que definir los deslizadores. Entonces el investigador solicitó a los participantes que observaran el comportamiento de los datos y después les preguntó, ¿el ángulo de inclinación de la recta que ajuste a los datos, es mayor o menor a 90º? Los alumnos al ver el diagrama de dispersión mencionan que el ángulo de inclinación de la recta debía ser mayor a 90º y en consecuencia el valor de la pendiente debía ser menor a cero, es decir negativo. A partir de esta discusión grupal, ellos ya tenían una idea de los valores que debían de darle al intervalo que controla el deslizador de la pendiente *m*.

Una situación similar sucedió con la ordenada al origen *b*, debido a que los participantes tuvieron dificultad de identificar tanto el intervalo como el incremento que debía dársele al deslizador que controlara dicho parámetro. El investigador les pidió a los alumnos que activen los ejes cartesianos en la vista gráfica de GeoGebra, para tener una referencia, después se les sugirió que realizaran un ajuste 1:1 en la escala de la vista gráfica y realizaran un zoom de alejamiento hasta visualizar ambos ejes coordenados para poder tener una referencia. Entonces los participantes ya pudieron visualizar el diagrama de dispersión y comentaron que luce muy pequeño en la parte baja de la izquierda muy pegada al eje X (Figura 35). Ya con esta vista los alumnos, pudieron tener una idea de que los valores de la ordenada al origen deben ser positivos y cercanos a 1000.

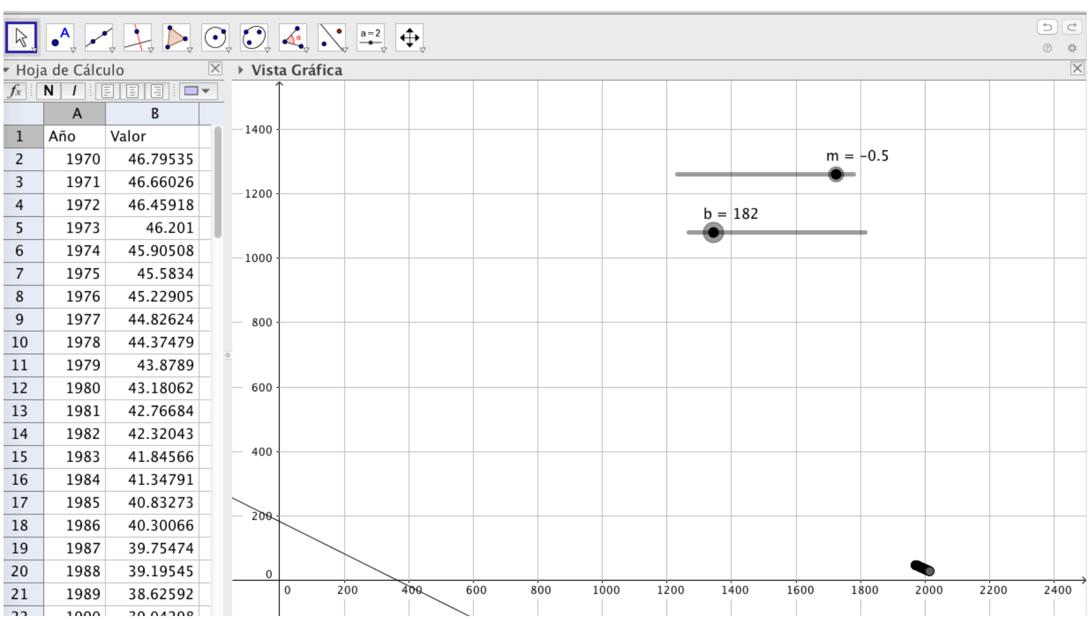


Figura 35. Vista del diagrama de dispersión con una escala 1:1

Ahora se les solicitó a los alumnos que movieran los deslizadores, para que pudieran conjeturar, cuáles son los valores de los parámetros *m* y *b* que generaban una línea recta que ajustaba con el diagrama de dispersión y en caso de ser necesario, ajustaran nuevamente esos valores ampliando o disminuyendo los intervalos y los incrementos de los deslizadores (Figura 36).

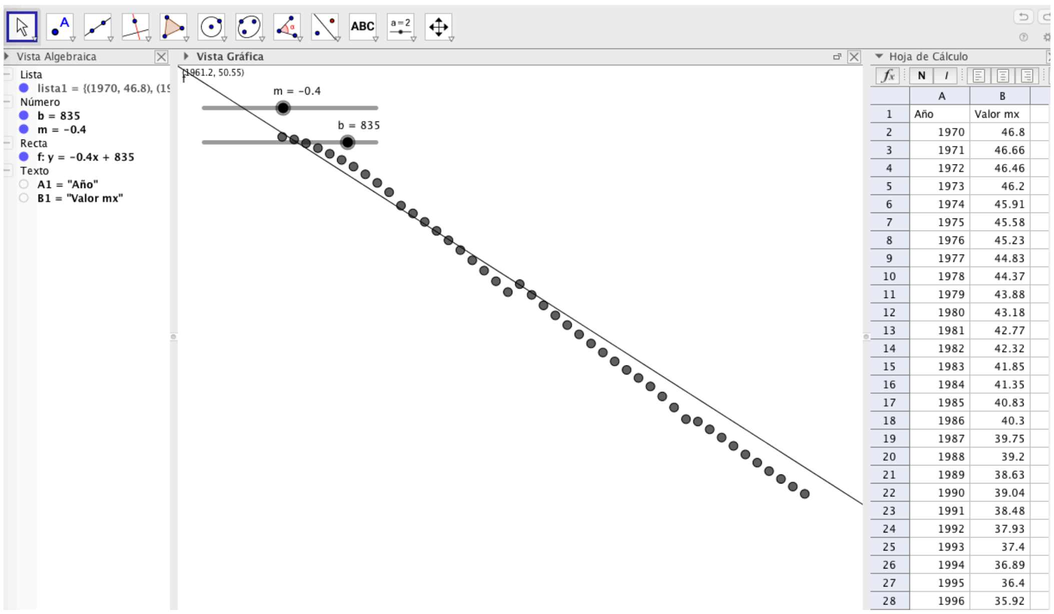


Figura 36. Vista del diagrama de dispersión con el ajuste de la recta hecho con los deslizadores

En la actividad aparecían una serie de preguntas que debían contestar los participantes, esto con el fin de que reportaran los acercamientos de la solución del problema, las justificaciones de las conjeturas, así como el significado que le dan a los resultados matemáticos obtenidos.

Las preguntas fueron:, ¿se parece o no tu resultado gráfico con el obtenido analíticamente?, ¿qué signo tiene el valor de la pendiente “m” en la recta ajustada?, ¿qué significado tiene el signo de “m” ?, ¿cuál es el valor de la población joven en México para el año 2020?, ¿en qué año aproximadamente el valor de la población juvenil en México será cero?, ¿cuál es un requisito indispensable para que suceda lo anteriormente cuestionado?, ¿consideras que la recta de regresión lineal obtenida es un buen ajuste?
Además, se les pidió que buscaran en internet el concepto de coeficiente de correlación, con la finalidad de contestaran las preguntas, ¿cuál es el rango para el coeficiente de correlación?,

¿cuál es el valor del coeficiente de correlación que se obtiene con GeoGebra del ajuste de la línea recta?, ¿consideras que la recta de regresión lineal obtenida en el inciso "h" es un buen ajuste?, ¿qué acciones o medidas en términos de políticas públicas podrías sugerir que se tomen en base a los resultados obtenidos?

Las respuestas que dieron todos los participantes a la pregunta, ¿por qué consideras que son un buen ajuste esos valores de "m" y "b"? se dieron en dos sentidos, Nisa Danahe, Laura Selene, Iván, Victoria, Adrián, Adriana, Fernanda y Ángel, reportan argumentos de cercanía de la línea que ellos proponen con los puntos del diagrama de dispersión (Figura 37).

Laura Selene:

Considero que es un buen ajuste con los valores de m y b, porque la recta se encuentra lo más cercana posible a los puntos del diagrama de dispersión.

Ángel:

Por qué la recta atraviesa por los puntos de dispersión, mostrando la dirección en la que van dichos puntos

Figura 37. Evidencias de Laura Selene y Ángel

Kenya, Daniela, Brenda, Alfonso y Aniloreny, además de dar sus argumentos reportan las conjeturas para realizar el ajuste (Figura 38)

Kenya:

Los valores que creo se ajustan más son; m=-0.5 y b=1033 porque son los valores en los que la recta y = m x + b se acerca a la mayor parte de los puntos del diagrama de dispersión.

Aniloreny:

Y=-0.42x+874

Para utilizar el deslizador de "m" primero observe que la gráfica de dispersión estaba inclinada hacia abajo por lo que entonces supuse que la pendiente tendría que ser menos de 0 por lo que entonces le di valores de mínimo -0.9 y de máximo 5.

Al deslizador de b "La ordenada al origen" le di valor de 900. Al activar la animación de los deslizadores observe que cuando la recta quedaba en Y=-0.42m+874, esta se acomodaba bien a los puntos de dispersión de la gráfica, de manera que se viera una recta como gráfica.

Figura 38. Evidencias de Kenya y Aniloreny

Después de que los participantes realizaron el ajuste de forma visual, se les pidió que con esos mismos datos de la población juvenil y el año correspondiente obtuvieran el ajuste de la recta empleando la opción de regresión lineal con la que cuenta GeoGebra. Todos los alumnos reportan correctamente el resultado es (y = – 0.45x + 931.36).

Enseguida se les solicitó a los alumnos que compararan los resultados de los ajustes visual y el hecho con la herramienta para que argumentaran su parecido o diferencia (Figura 39).

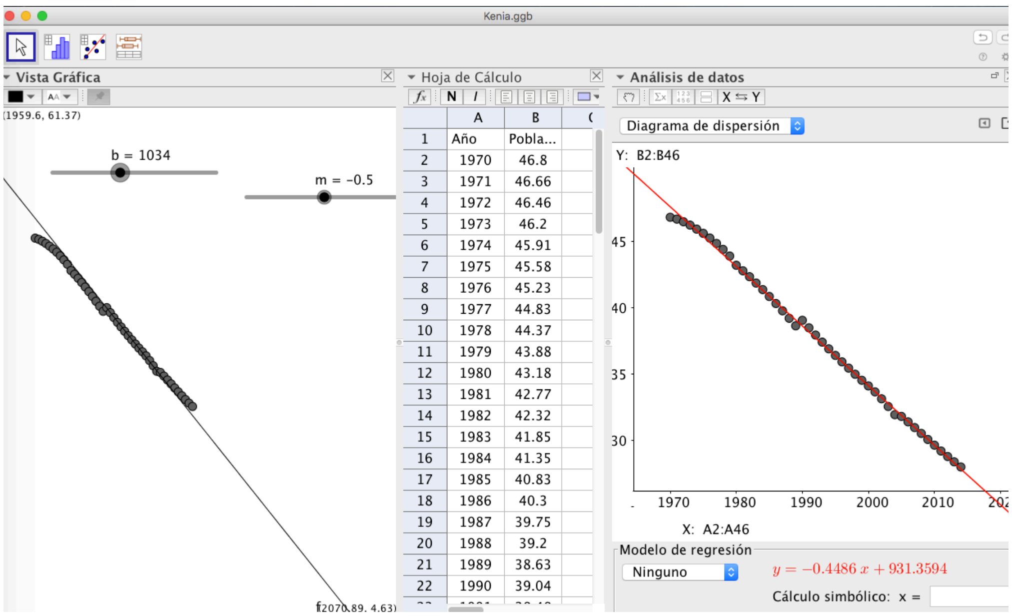


Figura 39. Comparativo de los ajustes visual y analítico

La intención de esta comparación, fue que los alumnos pudieran constatar si el ajuste hecho de forma visual era similar al hecho analíticamente. En las respuestas de los alumnos ofrecen una argumentación de la similitudes y diferencias. Kenya, Nisa Danahe, Iván, Victoria, Adriana, Alfonso, Aniloreny, Fernanda y Ángel, reportan que sus resultados son muy parecidos (Figura 40)

Alfonso:

En mi ejercicio me quedo

Línea recta supuesta

$\boxed{y = -0.45x + 924}$

En el ejercicio correcto quedo

Línea recta ajustada

$\boxed{y = -0.45x + 931.36}$

Fernanda:

Línea recta supuesta: y= -0.46x + 950

Línea recta ajustada y= -0.45x + 931.36

En la ecuación del inciso "h" es mucho más exacta y con ella los puntos quedan mejor ajustados y en el inciso "e" como está hecho al tanteo la ecuación hace que los untos no sean tan ajustados.

Figura 40. Evidencias de Alfonso y Fernanda

El resto de los participantes al contrastar sus resultados, argumentan que su ajuste hecho de manera visual no fue tan bueno (Figura 41)

Victoria:

Con deslizadores m = -0.46; b = 953.9
Con la herramienta m = -0.45; b = 931.36
En mi resultado aumentó .01 la pendiente y en cuanto la ordenada al origen el programa usó una menor por menos de veinte unidades, en cierto punto podrían parecer pequeños los cambios (en la manera en que se ve en la gráfica), pero en la realidad esas variaciones pueden dar como resultado malas predicciones, por ejemplo, mi resultado predice que el nivel de jóvenes decaerá antes debido a su mayor inclinación.

Aniloreny:

Línea recta supuesta Y=-0.42x+874

Línea recta ajustada Y=-0.45x+931.36

No se parecen, la ecuación de mi recta supuesta solo tiene una poca cantidad de diferencia con la línea ajustada, la recta ajustada se adapta mejor a la gráfica de dispersión representada.

Figura 41. Evidencias de Victoria y Aniloreny

Antes de continuar el investigador les comenta a los alumnos que si bien es cierto algunos de los ajustes hechos con los deslizadores son más parecidos que otros al resultado obtenido de manera analítica, es necesario considerar la cantidad de decimales que se usan al construir los deslizadores. Pero todos los resultados son una primera aproximación de un ajuste de línea recta.

El investigador cuestiona a los alumnos por el signo de la pendiente de la línea de ajuste, y todos los alumnos identifican que es negativo. Después contestaron la pregunta, ¿qué significado tiene el signo de *m* en términos de la población joven?

Kenya, Nisa Danahe, Daniela, Iván, Victoria, Adrián, Adriana, Alfonso si identifican que es la relación que existe entre la proporción de población joven con respecto al paso de los años y dado que el signo de la pendiente es negativo, esto da como resultado una disminución de la proporción de población juvenil con respecto al tiempo (Figura 42)

Nisa Danaeh:

Que mientras mas pasan los años, la poblacion joven se va reduciendo en México.

Ivan:

Significa que la población jovénes disminuye, cada vez son menor la población jóvenes

Figura 42. Evidencias de Nisa Danahe e Iván

Mientras que el resto de los participantes, tuvo dificultad para identificar el significado del signo negativo en el ajuste de la línea recta. Algunos reportaron que se trataba del porcentaje de la población, o valor más común (Figura 43)

Fernanda:

El porcentaje que hay de la población.

Brenda:

El promedio o valor más común entre este sector social. Posiblemente se refiere al número o cantidad mas constante /frecuente.

Figura 43. Evidencias de Fernanda y Brenda

En el siguiente inciso de la actividad se les pidió a los alumnos que calcularan con uso de la herramienta tecnológica, el valor de la población joven en México en el año 2020

Todos los alumnos reportaron con éxito el valor aproximado de 25.15, con pequeñas variaciones debido a los decimales. Brenda no supo emplear la herramienta para calcular el valor solicitado, pero basó su respuesta en la gráfica y su resultado en efecto como ella lo menciona, no es exacto, pero si se aproxima bastante (Figura 44).

Brenda:

No supe como poder obtener esta cantidad. Aunque como se pude observar, la cantidad de jóvenes va disminuyendo a través de los años así que puede ser una cantidad aproximada de 25.0000000.

Figura 44. Evidencia de Brenda

Posteriormente se les solicitó a los participantes que calcularan el año aproximadamente en el cual el valor de la población juvenil en nuestro país será cero, así como que mencionaran un requisito indispensable para que esto suceda.

En el caso de la obtención del año alumnos reportan que es aproximadamente en 2076, con pequeñas variaciones dependiendo de la cantidad de decimales con los que trabajaron en la herramienta. Kenya fue la única alumna que indica cómo se obtenía dicho resultado (Figura 45)

Kenya:

En el año 2069, ya que al igualar la ecuación a 0, y despejar la variable de x, da ese resultado.

$$x = \frac{-931.36}{-0.4486} = 2076.14$$

Figura 45. Evidencia de Kenya

La argumentación que dieron los alumnos a algún requisito para que en algún año futuro (2076), el índice de población joven en México sea cero, las respuestas van en dos vertientes, por un lado, Laura Selene, Brenda, Iván, Alfonso y Aniloreny, escriben argumentos, como que el comportamiento de la población joven siga el mismo patrón (Figura 46)

Brenda:

Que el patrón o el seguimiento sea similar a través de los años

Ivan:

Para que suceda que la población joven llegara a cero entre el 2076 y el 2077 es que la población juvenil continúe disminuyendo en forma que propone la ecuación dada por GeoGebra.

Figura 46. Evidencia de Kenya

Kenya, Nisa Danahe, Daniela, Victoria, Adrián, Adriana, Fernanda y Ángel mencionan que las familias deben tener menos hijos, y que con el paso de los años va a existir más ancianos y menos niños.

Para finalizar la actividad, se les solicita a los alumnos buscar en internet el concepto de coeficiente de correlación, para después calcular este valor y con base en la información que obtenida contesta si, ¿consideras que la recta de regresión lineal obtenida en el inciso “h” es un buen ajuste?

Antes de iniciar la búsqueda de la información en internet, el investigador les insistió a los alumnos que no registraran la primera información que encontraran, les sugirió que buscaran 2 o 3 sitios y que trataran de interpretarlo con sus propias palabras. Porque después tenían que encontrar el valor en GeoGebra y dar una interpretación de ese valor.

Las respuestas que reportan todos los alumnos son correctas, el valor del coeficiente de correlación para la línea recta ajustada es de –0.9992, algunos estudiantes como Iván y Aniloreny (Figura 47) ofrecen una interpretación acertada del parámetro encontrado.

Ivan:

El valor de coeficiente de correlación es de -0.9992 por lo cual existe una correlaciona negativa mientras los años aumentan la población joven disminuye

Aniloreny:

Sí, es bueno el ajuste, la gráfica de dispersión se acomoda bien a la recta.

Debido a que el coeficiente de correlación se acercó más en este caso a -1, la línea recta será negativa pero tendrá una muy alta confiabilidad de hacer pronósticos porque el coeficiente que resulto fue r= -0.9992 y esto se acerca mucho a r=-1

Figura 47. Evidencia de Iván y Aniloreny

4.3.1 Análisis de la tercera actividad

Los alumnos en un principio no tenían claro, cuál era la intención de buscar dependencias o instituciones, que tuvieran información poblacional (en nuestro caso), pero se discutió la importancia que tiene el conocer instancias dónde puede buscarse información, por ejemplo información económica, política, de salud, etc. Para realizar la busqueda de información, la internet es una herramienta digital preponderante, pero también se volvió a remarcar la necesidad de ser críticos con la información encontrada (Figura 32).

Al inicio de la actividad los alumnos evidencian facilidad al graficar los datos de la población joven de México con la ayuda de GeoGebra (Figura 34), también muestran que no tienen dificultad al identificar que la forma que siguen los datos de la población joven México tiene una apariencia de línea recta. Se discutió con los alumnos que la gráfica de dispersión ayuda a deducir la intensidad de la relación (a través de la mayor o menor dispersión de la nube de puntos), visualizar su sentido (si la relación es directa o inversa) y el tipo (lineal o no), observando su comportamiento (tendencia).

Una vez que la gráfica de dispersión está construida surgió una discusión debido a que tuvieron que identificar los signos de los parámetros *m* y *b*, con base al comportamiento que sigue la gráfica (Figura 36). Algunos alumnos tienen presente los significados matemáticos de la pendiente y la ordenada al origen, pero la mayoría no. Es notorio que los alumnos consideran que los conocimientos adquiridos de geometría, aritmética y álgebra poco o nada tienen que ver con la probabilidad y estadística. En esta parte de la actividad los deslizadores de la herramienta digital cobran relevancia, porque estos permiten el movimiento de los valores asignados a los parámetros en un rango determinado y al moverlos se puede observar el comportamiento de los objetos para realizar conjeturas.

En la sección que se les pide a los estudiantes encontrar la recta de regresión lineal con GeoGebra, no reportan ningún problema en el manejo del software, pero es de llamar la atención que al realizar la comparación el ajuste hecho de forma visual y con el ajuste analítico (Figura 39), algunos alumnos dicen que sus ajustes visuales no son correctos, aunque la variación entre ambos es muy poca, lo que pone de manifiesto que los alumnos no consideran correctas las aproximaciones o estimaciones que se realicen de un problema.

Después de que los alumnos ya cuentan con la recta de regresión lineal que ajusta a los valores de la población joven en México, se les solicitó que estimaran, el valor de la proporción para el año 2020 (Figura 44) y el año aproximado en el que este parámetro es cero aproximadamente (Figura 45), no reportan problemas. Esto se debe que GeoGebra permite hacer estos cálculos, permitiendo a los estudiantes dejar de lado la carga de operaciones aritméticas y algebraicas, para centrarse en encontrar una explicación y sentido a los resultados estadísticos obtenidos (Figuras 46).

Cuando se les solicitó a los alumnos que investigaran el concepto de coeficiente de correlación, calcular el valor del mismo y que juzgaran si el ajuste de regresión lineal que habían hecho de los valores de la población joven en México, las respuestas reflejan una comprensión del concepto, así como una interpretación adecuada del parámetro en cuestión (Figuras 47). El uso del internet en la búsqueda de información y GeoGebra como herramienta estadística son determinantes en el éxito de las respuestas que los alumnos proporcionan.

Capítulo 5. Conclusiones

En este capítulo se presentan las conclusiones del estudio. Se intenta dar respuesta a las preguntas de investigación. Estas preguntas tienen el objetivo de documentar y analizar las características que debe tener un ambiente de aprendizaje, en el cual se incorpora diversas herramientas tecnológicas y el uso de información real. Se muestran los acercamientos, formas, errores y resultados obtenidos por un grupo de alumnos de bachillerato cuando trabajan en un ambiente de resolución de problemas, en donde las preguntas son el medio para la comprensión de conceptos probabilísticos y estadísticos. Al final del trabajo se realizan algunas reflexiones y limitaciones con respecto a los resultados del estudio y que pueden dar pauta a trabajos de investigación futuros.

5.1 Conclusiones

Primera pregunta de investigación, ¿de qué manera la implementación de actividades en un ambiente de resolución de problemas y el uso coordinado de herramientas digitales contribuyen a la comprensión de un evento aleatorio en un grupo de alumnos de bachillerato?

El papel de las herramientas digitales

Los resultados muestran que GeoGebra puede ser usada como una herramienta de exploración que simula el evento de lanzar un par de dados y registrar la suma de las caras superiores. Esta herramienta fue proporcionada por el investigador para que los alumnos pudieran contrastar y verificar si el juicio intuitivo que siguieron al inventar una serie de datos que pudieran parecer el resultado del lanzamiento y suma de las caras superiores de un par de dados, se aproxima al obtenido si se toman en cuenta todos los pares ordenados generados por el espacio muestra o si, por el contrario, solo pensaron en el valor de la suma y no consideraron los diferentes pares que generan el valor de la suma de los dados. El deslizador de la aplicación controlaba el número de veces que se simula el lanzamiento de los dados y permitió potenciar las ideas matemáticas de los alumnos, ya que con la modificación del deslizador se modificaba el número de simulaciones, permitiendo observar el comportamiento que siguen los valores de las frecuencias al aumentar el número de experimentos. En este sentido, la simulación permite que los alumnos visualicen el comportamiento que sigue un fenómeno aleatorio si este se lleva a cabo un número considerable de veces, tarea que resultaría laboriosa haciendo uso de material manipulativo. En nuestro caso el lanzar 1000 veces un par de dados sería una tarea por demás complicada.

GeoGebra y Excel fueron las herramientas empleadas para graficar y explorar los datos generados por los alumnos en la actividad del lanzamiento de los dados. Es importante señalar que en el caso de Excel ofrece la facilidad de generar graficas de barras de una manera sencilla para los alumnos. Pero es importante mencionar que tiene la desventaja de que algunos diagramas como el de caja y bigotes es complicado su elaboración, por tal motivo se optó por usar GeoGebra como alternativa para realizar esta tarea, además de que este software también proporciona las medidas de tendencia central y de dispersión de las distribuciones estudiadas. Lo relevante es que los alumnos conozcan diferentes opciones de herramienta digitales para que cuando tengan que representar información de una forma gráfica tengan alternativas y usen la que más les convenga y sea útil.

Internet es una herramienta importante cuando los alumnos necesitan buscar información de conceptos estadísticos, como evento aleatorio o la distribución de probabilidad. Pero es evidente que los estudiantes están acostumbrados a quedarse con la primera información que el buscador les proporciona.

Por esa razón al solicitarles que hagan uso de esta herramienta digital es pertinente que el profesor promueva el pensamiento crítico para poder discernir la información valiosa y verídica de la que no lo es.

El papel del estudiante

Durante el desarrollo de las actividades analizó el desempeño de los estudiantes para identificar cómo fue evolucionando la percepción que ellos tenían de un evento aleatorio. En el transcurso de la actividad propuesta, los jóvenes realizaron conjeturas al manipular una aplicación en GeoGebra, variando el número de simulaciones, para que encontraran patrones de comportamiento de una distribución de frecuencias con la finalidad de contrastar sus pensamientos intuitivos al inventar una serie de datos supuestamente aleatorios. Además, tomaron decisiones propias para elegir el tipo de gráfica que ellos consideraban como la mejor opción para representar la información recopilada. Es importante mencionar que les cuesta trabajo reportar justificaciones de los resultados obtenidos, debido a que no están acostumbrados a ser reflexivos con los mismos. Y en el caso de la búsqueda de información en internet, es cierto que es una práctica común para ellos, pero se presenta de una forma desordenada y obedece a la inmediatez, no son autocríticos, por eso les resulta extraño que se les solicitara una interpretación o paráfrasis de lo encontrado además de citar las fuentes consultadas para una posible contrastación o verificación de la información.

El papel del profesor

Las participaciones del profesor fueron muy importantes, realizando una serie de preguntas y propuestas de autorreflexión, que guiaron a los alumnos durante el desarrollo de las actividades. Las intervenciones del profesor se dieron en el transcurso de las sesiones, verificando las respuestas de los alumnos e interviniendo de forma grupal. La finalidad de esas intervenciones es que los alumnos posean una idea del propósito que tienen las actividades.

La tarea de acompañamiento y seguimiento de los jóvenes es complicada y laboriosa, ya que de manera continua debe estar atento a las reacciones y respuestas de los alumnos, para solventar dudas o proponer situaciones que detonen las ideas estadísticas. En caso de ser necesario proponer observaciones del comportamiento de patrones, búsquedas de información, manipulaciones de objetos (dados) y autorreflexiones de los resultados obtenidos. Ese seguimiento presencial, además debe ser complementado con un análisis de la información que los participantes van registrando en la actividad, en ese sentido la plataforma Google Classroom es una excelente opción, porque el profesor tiene la oportunidad de ver las respuestas de los alumnos y en caso de ser necesario, realizar intervenciones, con la finalidad de hacer aclaraciones, corregir errores e identificar la evolución de los alumnos.

Segunda pregunta de investigación, ¿cómo influye el uso de las herramientas digitales y la utilización de datos reales en una clase de probabilidad y estadística para realizar inferencias matemáticas?

Ben-Zvi y Bakar (2016), señalan que el creciente interés por la comprensión de la estadística, ha llevado a un cambio de paradigma en la conceptualización de la enseñanza estadística y el desarrollo de los profesores en esta área de la matemática. Además, sugieren modificaciones en los enfoques de enseñanza, reflexionando sobre el papel y el propósito de las estadísticas. Para lograrlo proponen que los maestros diseñen e implementen cursos donde se involucren datos reales e información estadística que esté al alcance de los alumnos.

El papel de las herramientas digitales

En la actividad del ajuste de recta, GeoGebra también fue una muy buena alternativa para graficar los datos y explorar el comportamiento de los mismos. Al realizar el diagrama de dispersión con la ayuda de la opción *análisis multivariable*, se da la oportunidad de que los alumnos observen el comportamiento de los datos. En la gráfica se aprecia que los puntos se distribuyen alrededor de una recta y entonces se procede a realizar un análisis de regresión lineal.

El uso de los deslizadores, potencio el pensamiento estadístico de los alumnos, ya que pudieron realizar conjeturas de los valores de la pendiente m y la ordenada al origen b de la línea recta que ajustaba mejor a los datos. Además, este software cuenta con herramientas para determinar el ajuste de curvas y proporciona la información estadística del ajuste, lo cual trae como consecuencia una disminución considerable de las operaciones numéricas y se puede hacer énfasis en la interpretación y valoración de los resultados estadísticos. El internet nuevamente formo parte importante como herramienta de búsqueda de información, para que los alumnos indagaran el concepto del coeficiente de correlación o para encontrar instituciones que ofrezcan información estadística de nuestro interés.

La importancia de los datos reales

En nuestros días el avance tecnológico ha impactado en la mayoría de los campos del conocimiento donde se realizan trabajos y en los cuales se efectúan mediciones, observaciones o experimentos de donde se obtengan datos de diferentes variables; es fundamental determinar algún tipo de relación de dependencia entre las variables con el fin de hacer predicciones o pronósticos de eventos futuros de acuerdo con el comportamiento de ellas. El auge del internet, ha facilitado que las personas tengan la posibilidad de tomar datos e información de dependencias de gobierno, organizaciones no gubernamentales, empresas y demás instituciones, con la intención de realizar un análisis para determinar cómo lo conocido se relaciona con el evento futuro y así ayudar a determinar una solución eficaz de los problemas que se presentan. Un procedimiento estadístico que se utiliza para este fin se conoce como análisis de regresión que permite establecer la relación funcional o ecuación matemática que relaciona las variables, así como la fuerza de esa relación.

En este trabajo, el investigador les proporcionó a los participantes información obtenida de la OCDE, el indicador seleccionado fue la población joven. Este organismo define a la población joven como las personas que pueden depender de la ayuda de otros para sus necesidades diarias con respecto al número de aquellas personas que son capaces de proporcionar ese apoyo. Es llamativo que la población joven en México y el tiempo en los últimos 44 años se vinculan de una forma que su representación gráfica aparenta una relación lineal inversa, es decir que sigue la forma de una línea recta con pendiente negativa.

La importancia de emplear datos reales en una clase de matemáticas con jóvenes de bachillerato radica en que sean capaces de constatar la utilidad e importancia de la de estadística. Además de que, tomando información real, se les puede motivar a que se sientan como si fueran profesionales de la estadística y propongan interpretaciones de los resultados inferidos.

5.2 Reflexiones finales

Las actividades propuestas fueron el eje esta investigación, buscando que, en un ambiente de resolución de problemas, los estudiantes de bachillerato mejoraran sus percepciones del azar, y de esta forma pudieran comprender lo que es un evento aleatorio. Los alumnos tuvieron oportunidad de generar los datos con los que se trabajó en la actividad, graficar esos datos, compararlos y obtener medidas de tendencia central. Fue relevante el uso de la tecnología, en este caso el deslizador que controlaba el número de veces que se simulaba el experimento fue la herramienta que permito potenciar el pensamiento de los jóvenes, ya que pudieron observar el comportamiento de la distribución de frecuencias del lanzamiento de los dados, cuando se aumenta el número de simulaciones, de esta manera pudieron conjeturar que esa distribución es muy parecida a la distribución de probabilidad si el número de experimentaciones es grande.

La primera y segunda actividad permitieron crear además el vínculo entre la estadística y la probabilidad; porque los jóvenes recolectaron, graficaron y analizaron información, que es precisamente algunas de las actividades de las que se ocupa la estadística. Pero esa información provenía de un experimento aleatorio, y es ahí donde entra el estudio de la probabilidad. Una vez que los alumnos tengan un acercamiento a la probabilidad con el lanzamiento de un par de dados, se puede extender a otros eventos aleatorios, en donde ellos con la forma que se les propuso trabajar, empleen sus conocimientos (recursos), puedan realizar conjeturas, construyan modelos de simulación del experimento, reporten resultados y modifiquen sus sistemas de creencias, con respecto al azar.

En la segunda actividad se trató un tema de datos bivariados, el cual en bachillerato muchas veces pasa desapercibido y que en algunos planes de estudio se encuentra ausente, pero es una herramienta estadística muy importante. Las personas siempre debemos tomar decisiones, las cuales pueden ser sencillas o muy complejas, y surgen de eventos determinantes en los cuales solo hay una solución y eventos en donde existen varias soluciones y que además no se tiene la certeza de que vaya a ocurrir. Es por esta razón que es una muy buena idea que los estudiantes conozcan, cómo realizar el ajuste de curvas. Los valores con los que trabajaron en esta investigación siguen un comportamiento lineal, y la intención era precisamente que los alumnos tuvieran un acercamiento con información real, pero que fuera además significativamente muy ajustado a una línea recta. Este puede ser un buen inicio para adentrar a los alumnos, para después dar paso a otros conceptos estadísticos, como la prueba de bondad de ajuste, la covarianza, o la potencia de una prueba.

Esta investigación puede contribuir a que la probabilidad y estadística sea enseñada y aprendida de una forma más dinámica, intuitiva y razonada. Si bien es cierto que esta rama de la matemática es obligatoria en nuestro país, la formación de docentes especialistas en ella es escaza, hace falta capacitación para futuros profesores y actualización para los que actualmente se encuentran frente a grupo. El objetivo sería generar una comunidad que entienda y promueva el aprendizaje y enseñanza de la probabilidad y estadística, y se discuta y promueva el uso de las tecnologías como algo habitual en su enseñanza.

Para finalizar se dejan abiertas algunas preguntas que podrían ser abordadas en investigaciones futuras, ¿cómo el uso coordinado de las tecnologías digitales puede contribuir a resolver problemas de probabilidad en un ambiente de resolución de problemas?, ¿cuáles son los recursos y heurísticas que un profesor de probabilidad y estadística debe evidenciar?

Revisión bibliográfica

Bakker, A., & Gravemeijer, K. P.E. (2004). Learning to reason about distributions. En D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147–168). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bakker, A., Kent, P., Noss, R., & Hoyles, C. (2009). Alternative representations of statistical measures in computer tools to promote communication between employees in automotive manufacturing. *Technology Innovations in Statistics Education*, 3(2), Article 1. Retrieved from http:// escholarship.org/uc/item/S3b9122r.

Batanero, C. (2014). Probability Teaching and Learning. En S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (pp 491-496). New York: Springer.

Begg, A. (1997). Some emerging influences underpining assessment in statistics. En I. Gal, y J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 17-26). Amsterdam: IOS Press.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning*, 2(1–2), 127–155.

Ben-Zvi, D. (2004a). Reasoning about data analysis. En D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 121–145). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ben-Zvi, D., Gil, E., & Apel, N. (2007, August). What is hidden beyond the data? Helping young students to reason and argue about some wider universe. In D. Pratt & J. Ainley (Eds.), *Reasoning about informal inferential statistical reasoning: A collection of current research studies*. Proceedings of the fifth international research forum on statistical reasoning, thinking, and literacy (SRTL-5), University of Warwick, UK, August, 2007.

Ben-Zvi, D. (2014). Data Handling and Statistics Teaching and Learning. En S. Lerman (Eds), *Encyclopedia of Mathematics Education*, (pp. 137 – 140). Dordrecht: Springer

Ben-Zvi, D. & Bakar K. (2016). International Perspectives on the Teaching and Learning of Statistics. En D. Ben-Zvi & K. Makar (Eds), *The Teaching and Learning of Statistics International Perspectives* (pp. 1 – 12). Switzerland: Springer.

Biehler, R. (1997a). Software for learning and doing statistics. International Statistical Review, 65, 167–189.

Biehler, R., Ben-Zvi, D., Bakker, A., & Makar, K. (2013). Technological forn enhancing statistical reasoning at the school level. En M. A. (Ken) Clements et al. (Eds.), *Third International Handbook of Mathematics Education* (pp. 643–689). New York: Springer.

Borba, M. C., Clarkson, P., & Gadanidis, G. (2013). Learning with the use of the internet. En M. A. (Ken) Clements, A. Bishop, C. Keitel, J. Kilpatrick & F. K. S. Leung (Eds.), *Third International Handbook of Mathematics Education* (pp. 691-720). New York: Springer.

Burril, G. & Biehler R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill, and C. Reading (eds.), *Teaching Statistics in School* 57 *Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study*.

Chance, B. L., & Rossman, A. J. (2006). Using simulation to teach and learn statistics. En A. J. Rossman & B. L. Chance (Eds.), *Proceedings of the seventh international conference on the teaching of statistics*. (ICOTS-7), Salvador, Bahia, Brazil, [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/ ~ iase/publications/17/7E1 CHAN.pdf

Chance, B. L., delMas, R., & Garfield, J. (2004). Reasoning about sampling distributions. En D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 295–323). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Collins (2003). *Collins English dictionary* (6Rev Ed edition). Collins.

Conapo (2013), *Proyecciones de Población 2010-2050, México*, En: http://www.conapo.gob.mx

Corominas (1973). *Breve diccionario etimológico de la lengua castellana*. Madrid, España: Gredos.

delMas, R. C., Garfield, J., & Ooms, A. (2005, July). Using assessment items to study students' difficulty with reading and interpreting graphical representations of distributions. Presented at the fourth international research forum on statistical reasoning, thinking, and literacy (SRTL-4), July 6, 2005. Auckland, New Zealand.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124–158.

Garfield, J., & Ben-Zvi, D. (2004). Research on statistical literacy, reasoning, and thinking: issues, challenges, and implications. En D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 397–409). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Garfield, J., & Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review / Revue Internationale de Statistique*. Vol. 75, No.3. (pp 372-396), International Statistical Institute (ISI)

Garfield, J., & Ben-Zvi, D. (2008). Developing Students' Statistical Reasoning. *Connecting Research and Teaching Practice*. New York: Springer

Gigerenzer, G. (2002). *Reckoning with risk*. Londres: Penguin Books.

Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la investigación*. México: McGraw W-Hill.

Kahnemann, D & Tversky, A. (1973). Avalibility: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207 – 232.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education*, 33(4), 259–289.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Satistics Education*, 2(1), 1–37.

Leung, F. (2013). Part III, Introduction to section C: Technology in the mathematics curriculum. En M. Clements, A. Bishop, C. Keitel, J. Kilpatrick, & F. Leung (Eds.) *Third International Handbook of Mathematics Education* (pp. 517-524). New York: Springer.

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, 23, 557-568.

Mills, J. D. (2004). Learning abstract statistics concepts using simulation. *Educational Research Quarterly*, 28(4), 18–33.

Moore, D. S. (2004). *The basic practice of statistics* (3rd ed.). New York: W. H. Freeman.

National Council of Teacher of Mathematics (2008). *The role of technology in the teaching and learning of mathematics. A position of the National Council of Teachers of Mathematics*. Recuperado el 21 de noviembre de 2016 de www.nctm.org/about/content.aspx?id=14233

National Council of Teacher of Mathematics (2011). *Focus in High School Mathematics: Technology to Support Reasoning and Sense Making*. Reston, VA: National Council of Teacher of Mathematics.

Pfannkuch, M., & Reading, C. (2006, November). Reasoning about distribution: A complex process. *Statistical Education Research Journal*, 5(2), 4–9. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/~ iase/serj/SERJ5(2).pdf

Polya, G. (1945). *How to Solve it*. Princeton: Princeton University Press.

Pratt, D. (2007). How do teachers foster students' understanding of probability? In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 182–201). The Netherlands: Kluwer Academic Publishers.

Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. En D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 201–226). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Santos-Trigo, M. (2007). La educación matemática, resolución de problemas, y el empleo de herramientas computacionales. Cuadernos de investigación y Formación en Educación Matemática, 8 (6), pp. 35-54.

Santos-Trigo, M., & Camacho-Machin, M. (2009). Towards the Construction of a Framework to Deal with Routine Problems to Foster Mathematical Inquiry. *Problems resources, and issues in mathematics undergraduate studies*. 19:3, 260-279, DOI: 10.1080/10511970701641990

Santos-Trigo, M. (2010a). A mathematical problem-solving approach to identify and explore instructional routes based on the use of computational tools. En J. Yamamoto, J. Kush, R. Lombard, & J. Hertzog (Eds.), *Technology Implementation and Teacher Education: Reflective Models* (pp. 208-313). IGI Global: Hershey PA.

Santos-Trigo, M. (2014). *La Resolución de Problemas Matemáticos: Fundamentos Cognitivos*. México D. F., México: Trillas.

Santos-Trigo, M., & Reyes-Martínez, I. (2014). The coordinate use of digital technologies in learning environments. En L. Uden, J. Sinclair, Y. Tao & D. Liberona (Eds.), *Learning Technology for Education in Cloud* (pp. 61-71). Switzerland: Springer.

Santos-Trigo, M. (2015). Sobre el hábito de preguntar. *Revista C2 Ciencia y Cultura* [en línea]. Recuperado de http://www.revistac2.com/sobre-el-habito-de-preguntar/

Santos-Trigo, M. (2016). Tecnologías digitales y educación. *Revista C2 Ciencia y Cultura* [en línea]. Recuperado de http://www.revistac2.com/tecnologias-digitales-y-educacion/

Santos-Trigo, M. (2017). Transición digital: ¿dónde estamos?, a ¿dónde vamos? *Revista C2 Ciencia y Cultura* [en línea]. Recuperado de http://www.revistac2.com/transicion-digital-donde-estamos-a-donde-vamos/

Shaughnessy J. M., Research in probability and statistics: Reflections and Directions En Grows D.A., (Editor), 1992, *Handbook of Research on Mathematics Teaching and Learning*, New York: Macmillan Publishing Company. pp. 465 – 494.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *The second handbook of research on mathematics* (pp. 957–1010). Information Age Pub Inc.

Schoenfeld, A. (1985). *Mathematical Problem Solving*. New York: Academic Press.

Wild, C. J. (2006, November). The concept of distribution. *Statistics Education Research Journal,* 5(2), 10–26. Retrieved July 15, 2007, from http://www.stat.auckland.ac.nz/~ iase/serj/ SERJ5(2) Wild.pdf

Zimmerman, C. (2005). The development of scientific reasoning: What psychologists contribute to an understanding of elementary science learning. Paper commissioned by the National Academies of Science (National Research Council's Board of Science Education, Consensus Study on Learning Science, Kindergarten through Eighth Grade). Retrieved April 26, 2008, from http://www7.nationalacademies.org/bose/Corinne Zimmerman Final Paper.pdf

Anexo 1. Actividades de la investigación

Primera Actividad – Cuestionario sobre el evento aleatorio

Nombre(s):

Temas considerados:
Evento aleatorio, fenomeno aleatorio y ludopatía.

Instrucciones:
Responde las siguientes preguntas con tus propias palabras, escribe lo más claro posible en este documento y súbelo al Classroom

Preguntas iniciales:

a) ¿Qué entiendes por un evento aleatorio?

b) Ejemplifica por lo menos tres fenómenos aleatorios:

c) En tú vida cotidiana consideras que existen fenomenos aleatorios, si es así mencionalos:

d) ¿De qué forma reaccionas al enfrentarte a un evento aleatorio?

e) ¿Qué importancia tiene el educar a los alumnos en temas que impliquen fenómenos aleatorios?

f) ¿Conoces qué es la *ludopatía*?, (si desconoces el término) buscalo en internet y con tus propias palabras escribe una definición

g) ¿Qué tan importante es que la población en México reconozco a la ludopatía, como una enfermedad?

h) ¿Conoces juegos de azar que influyan para que una persona desarrolle la ludopatía?

i) ¿Crees que los conocimientos acerca de probabilidad y estística influyan, para que una persona sea menos propensa a ser ludopata?

Segunda actividad – Lanzamiento de dados

Nombre:

Instrucciones:
De forma individual realiza las siguientes actividades y contesta cada pregunta, lo más claro posible.

Objetivo:
Se realiza un experimento para corroborar si los alumnos de bachillerato tienen buenas intuiciones respecto a los experimentos aleatorios. Además se contrastan los resultados obtenidos de manera experimental con los obtenidos si se inventan resultados y al realizar la simulación del mismo experimento con una herramienta digital.
El experimento consiste en lanzar un par de dados legales y sumar las caras superiores obtenidas.
La finalidad principal es hacer reflexionar al alumno sobre el hecho de que nuestras intuiciones sobre el azar en ocasiones nos correctas, para ello los alumnos realizarán la recolección de la información y haran un análisis de datos empleando la estadística, y encontrar la relación que existe entre la distribucón de frecuencias y la distribución de probabilidad del evento aleatorio.

Temas considerados:
Variable, recopilación de datos, tablas de distribución de frecuencias, representaciones gráficas (histogramas, polígonos de frecuencias, gráfica de barras, gráfica circular); Medidas de tendencia central, dispersión y de posición. Distrubución de probabilidad y esperanza matemáticas.

Los datos:
Los datos van a ser generados como resultado del experimento que será realizado por cada uno de los alumnos de la clase.

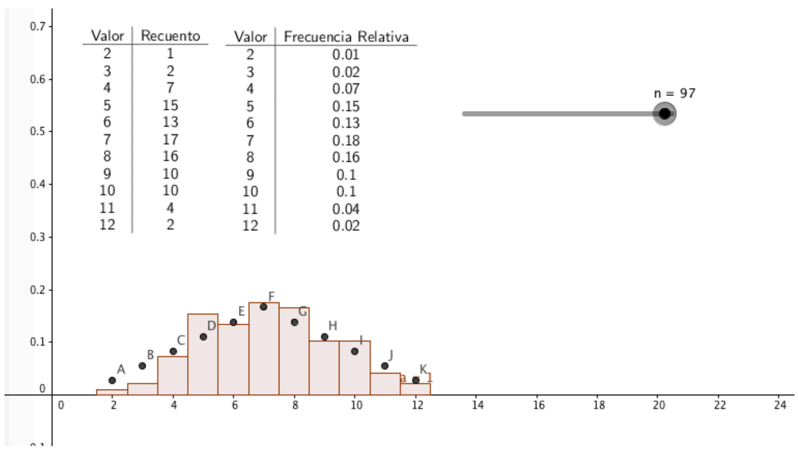
Herramientas tecnológicas:
Aplicación en Geogebra (simulador de la suma de las caras superiores del lanzamiento de un dado), hoja de cálculo.

Realización del experimento – El lanzamiento de un par de dados:
El experimento consiste en lanzar 20 veces un par de dados distinguibles (de diferentes colores), sumar el valor de los dados y registrarlo; después debes inventar una secuencia de 20 pares de caras posibles al lanzar el par de dados (sin lanzarlos realmente) de modo que la secuencia pueda pasar como aleatoria para otra persona y finalmente compararlos al realizar la simulación del lanzamiento y sumar las caras superiores de dos dados 20 veces con la ayuda de una aplicación hecha en Geogebra.

Registro de resultados:
En la siguiente tabla registra tus resultados obtenidos al lanzar realmente los dados y sumar las caras superiores.

En las siguiente tabla registra la suma de las caras de los dados tus resultados inventados.

En la última tabla registra tus resultados obtenidos en la simulación del lanzamiento de los dados y suma de sus caras superiores.



Valor	Recuento
2	1
3	2
4	7
5	15
6	13
7	17
8	16
9	10
10	10
11	4
12	2

Valor	Frecuencia Relativa
2	0.01
3	0.02
4	0.07
5	0.15
6	0.13
7	0.18
8	0.16
9	0.1
10	0.1
11	0.04
12	0.02

Mueve el deslizador "n", el cual representa el número de veces que se simula el lanzamiento de los dados y registra la suma de las caras superiores.

Preguntas en clase:

a) ¿Consideras que se puede distinguir una secuencia realmente aleatoria de otra que hemos inventado?

b) ¿Los resultados obtenidos son los mismos al realizar el lanzamiento de los dados, al inventar los datos, o al simular el experimento con una aplicación?

c) ¿Ahora que sucede si en la aplicación el experimento se simula 1000 veces?

d) ¿Tus resultados inventados se parecen a los obtenidos en el experimento o en la simulación?

El profesor interviene preguntando y registrando en el pizarrón, ¿qué te parece si comparamos el número de la suma de los dados en la secuencia real, la inventada y las obtenidas en la simulación de todos los alumnos de la clase?
Registro del número de la suma de los dados reales:

Número de suma al lanzar realmente los 2 dados	
Suma de puntos	Frecuencia
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

Registro del número de la suma de los dados inventados:

Número de la suma al inventar los valores de los 2 dados	
Suma de caras	Frecuencia
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

Registro del número de la suma de los dados obtenidos en la simulación:

Número de suma al simular el lanzamiento de los 2 dados	
Suma de caras	Frecuencia
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

Exporta cada una de las tablas hacia Excel y teniendo seleccionada la tabla con los datos de número de soles y frecuencia, en el menú principal escoge la opción de “insertar”, ahí escoge alguno de los gráficos recomendados: columna agrupada, barra agrupada, dispersión con líneas rectas y marcadores y línea.

a) ¿El gráfico que escogiste te permite comprender y resumir los datos de las suma de las caras superiores obtenidos en los tres casos (experimento, datos inventados y simulación)?

b) ¿Cuál es el valor más frecuente en cada uno de los casos?

Comparación de gráficos
Compara los tres gráficos y responde las siguientes preguntas:

a) ¿En qué se parecen?
b) ¿En qué se diferencian?
c) ¿El valor más frecuente es el mismo?
d) ¿Consideras que la forma en que pensaste los números al inventar los valores como si lanzaras 20 veces un par de dados legales es totalmente correcta?

Ahora construye una tabla donde concentres la información del experimento del lanzamiento real, inventado y simulado. Posteriormente esa tabla exportala nuevamente a Excel, selecciona los resultados de las tres tablas y nuevamente en el menú elige la opción de insertar y seleccona alguno de los gráficos recomendados donde puedas ver el comportamiento de las frecuencias de las tres tablas trabajadas.

Suma de caras superiores	Frecuencia real	Frecuencia inventada	Frecuencia simulada
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

Finalmente, en Geogebra realiza un análisis multivariable:

Observa la comparación de los tres gráficos y responde las siguientes preguntas:

a) ¿Tienen alguna misma medida de tendencia central (media, mediana, moda)?

b) Si el número de simulaciones en la aplicación en Geogebra fuese por ejemplo n = 1000, ¿cuántas veces crees que se vayan a obtener el número 7?

c) Si lanzas un par de dados distinguibles (diferente color) una sola vez, ¿cuántas veces puede aparecer 7 y cuáles serían los pares ordenados?

d) En terminos de probabilidad, ¿cuál será la probabilidad para que puede aparecer un 7 al realizar un solo lanzamiento el par de dados?, y ¿si simulas 1000 veces el lanzamiento del par de dados?

e) Investiga en internet el conceptos:
   - ✓ Distribución de probabilidad

f) Observa y compara la distribución de frecuencias con el deslizador en n = 1000 y la distribución de probabilidad del lanzamiento de un par de dados legales. Menciona si se parecen o difieren, ¿qué puedes decir al respecto?

Tercera actividad – Ajuste de una recta

Nombre:

Instrucciones:
De forma individual realiza las siguientes actividades y contesta cada pregunta, lo más claro posible.

Objetivo:
Realizar la recolección de información desde 1970 hasta 2014 de la población joven en el portal de internet de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), esto con el fin de infererir el comportamiento de este indicador mediante la obtención de una recta de regresión lineal.
La finalidad de la actividad es hacer reflexionar al alumno sobre el hecho de que mediante el uso adecuado de herramientas estadísticas como la regresión lineal y la correlación, se pueden inferir resultados futuros de una colección de datos históricos de una característica poblacional, describiendo la naturaleza e intensidad de las relaciones entre esos rubros mecionados y el tiempo.

Temas considerados:
Regresión lineal y correlación.

Los datos:
Los datos son obtienidos por los alumnos del portal de internet de la Organización para la Cooperación y el Desarrollo Económicos (OCDE):
https://data.oecd.org/pop/young-population.htm#indicator-chart

OECD (2016), Young population (indicator). doi: 10.1787/3d774f19-en

Características de la información a estudiar:

La población joven
La población juvenil se define como aquellas personas menores de 15 años. La proporción de la población dependiente se calcula como el total de la población de edad avanzada y la juventud expresa como una proporción de la población total. La relación de dependencia juvenil se refiere al número de personas jóvenes que pueden depender de la ayuda de otros para sus necesidades diarias con respecto al número de aquellos que son capaces de proporcionar ese apoyo. Las tendencias demográficas tienen una serie de implicaciones para el gobierno y el gasto privado en pensiones, asistencia sanitaria y la educación, y, más en general, para el crecimiento económico y el bienestar. Este indicador se mide como un porcentaje de la población. OECD (2016), Young population (indicator)

En tu sesión de Classroom, descarga y abre el archivo de Excel llamado Población juvenil, el cual fue obtenido de la página https://data.oecd.org/mexico.htm y que contiene la población joven (en términos porcentuales del total).

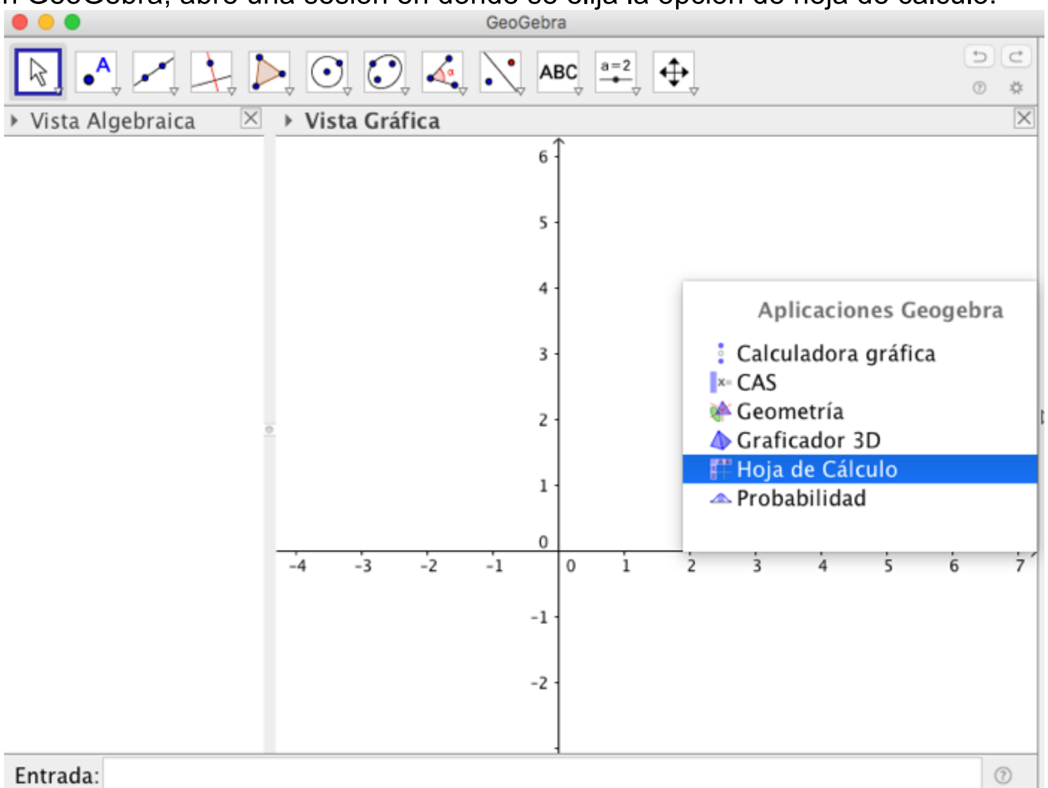
La información del indicador de población joven en México se muestra desde 1970 hasta 2014, la cual vas a encontrar en los renglones del 1008 al 1052.

Tu únicamente vas a trabajar con dos columnas (Time – Año) y (Value – Valor), estos datos seleccionales para cópiarlos en una tabla que vas a usar en Geogebra

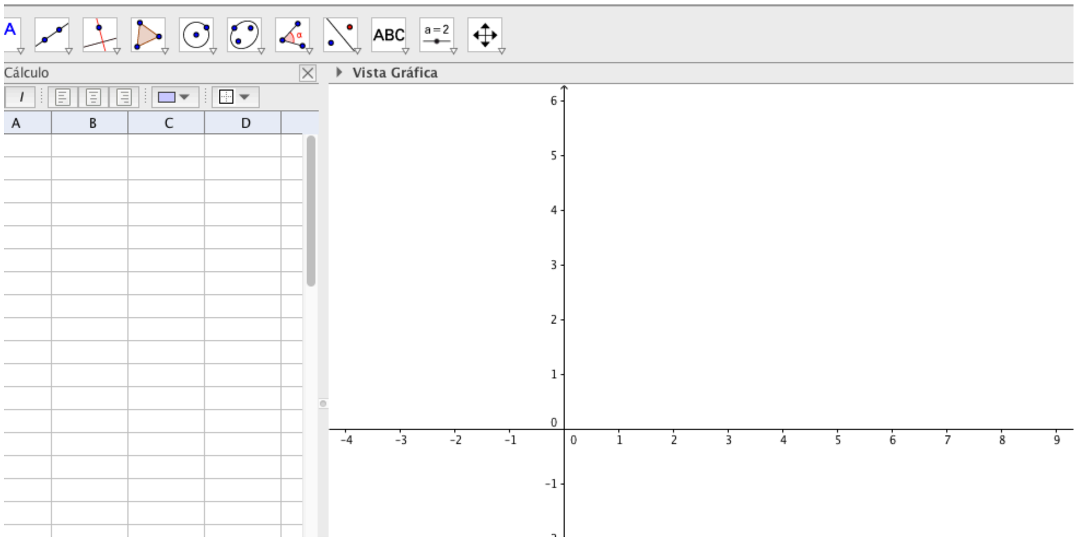
Año	Valor MX

Actividad:
Con los datos mostrados:

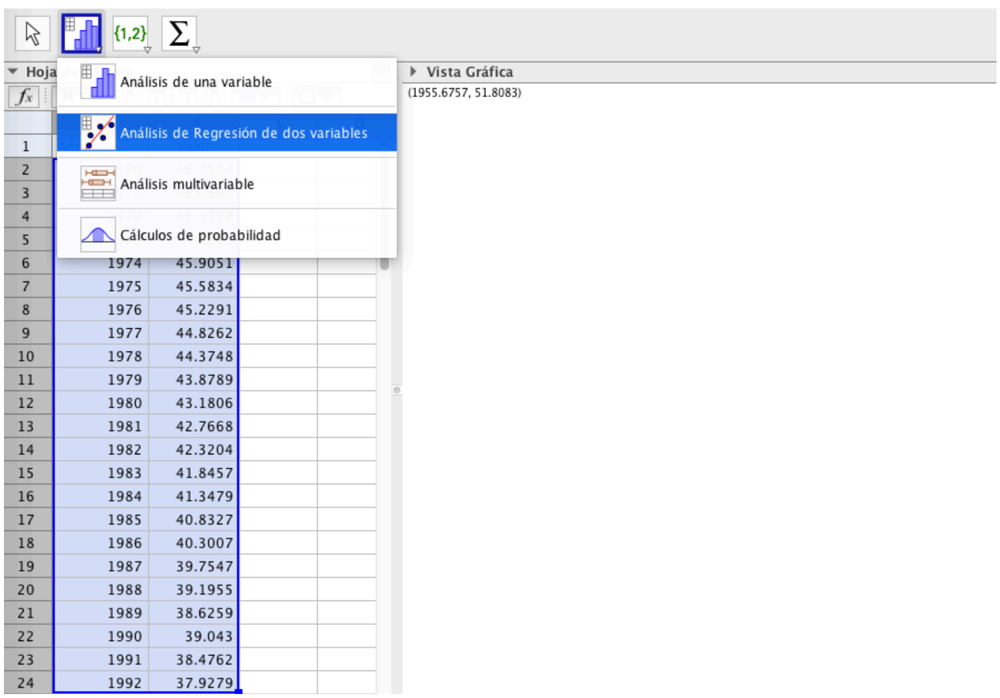
a) En GeoGebra, abre una sesión en donde se elija la opción de hoja de cálculo.



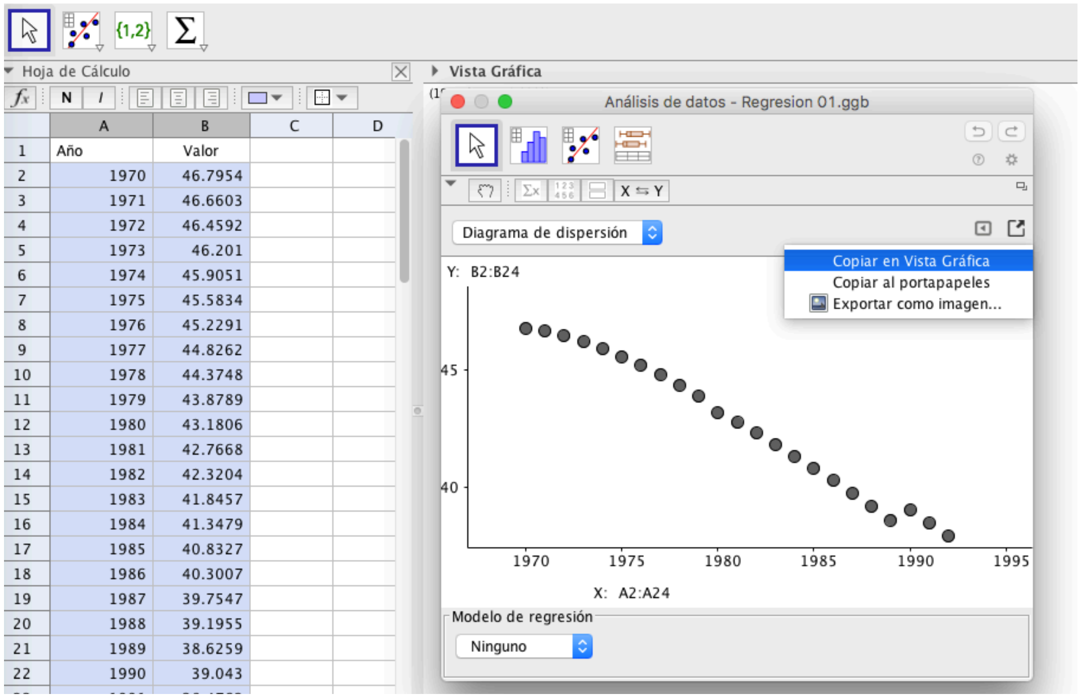
b) Cuando des clic a la opción de la Hoja de Cálculo, veras una pantalla como la siguiente:



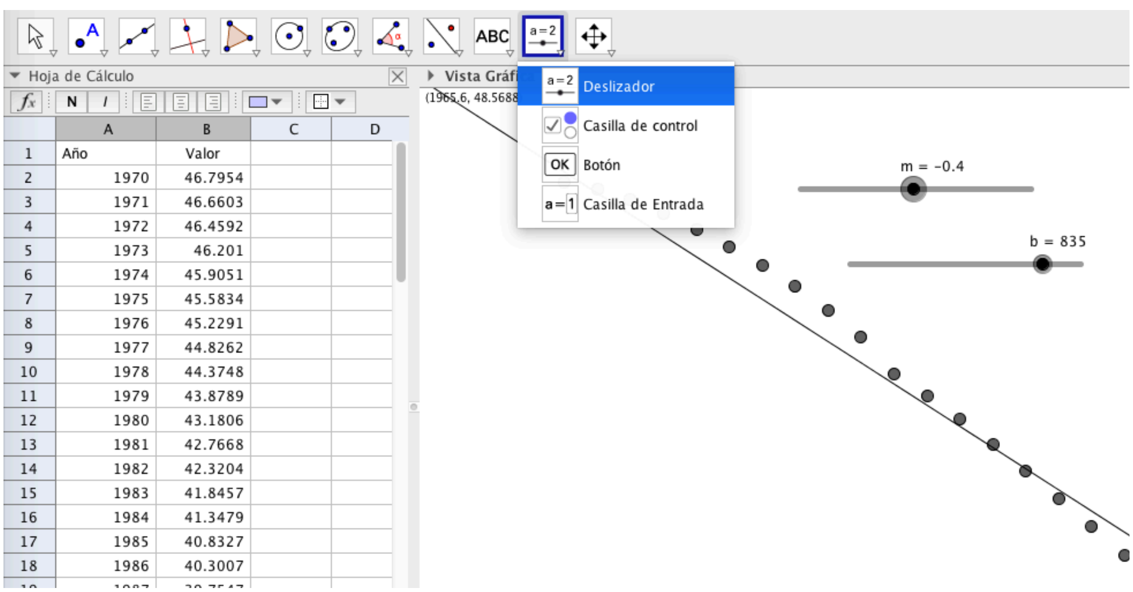
c) Introduce los datos a GeoGebra colocando el año en la columna A (el eje X) y el valor de la población jovén en México en la columna B (el eje Y), selecciona ambas columnas e inmediatamente se desplegarán varias opciones en el menú superior



d) Elije la opción "Análisis de Regresión de dos variables" y realiza una gráfica de dispersión, exporta la gráfica de dispersión a la vista gráfica.

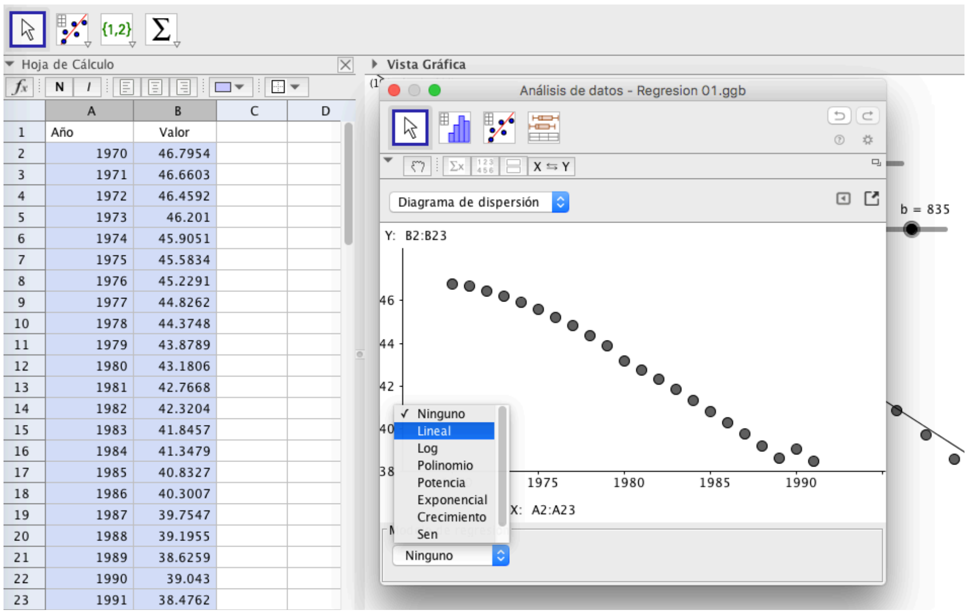


e) Crea dos los deslizadores, uno que sea de la pendiente “m” y el otro de la ordenada al origen “b”, busca de manera gráfica el mejor ajuste y registra los parámetros que sirven como coeficientes de la ecuación de la línea recta, a la cual vamos a llamar línea recta supuesta (y = mx + b). Anota los valores de “m” y “b” que consideres que se ajustan mejor al diagrama de dispersión.

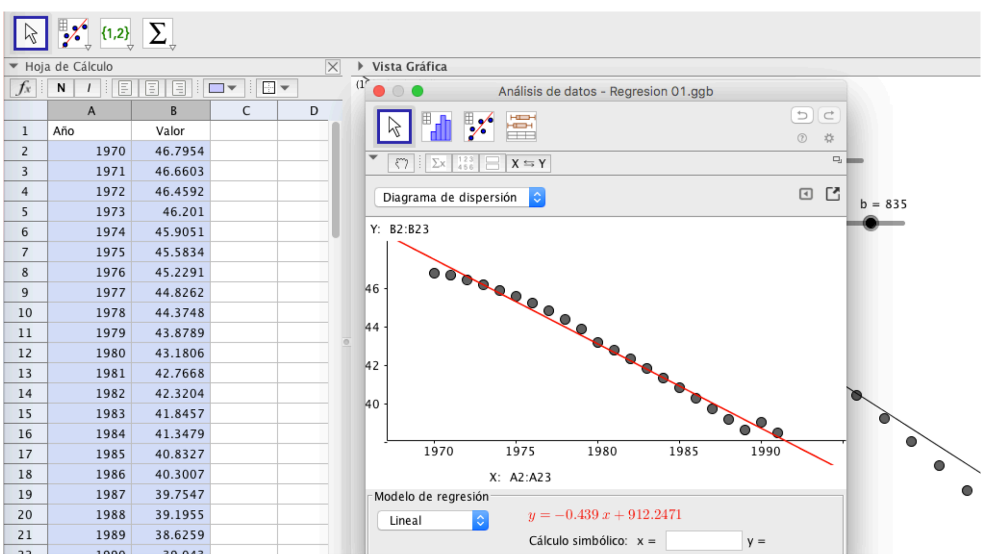


f) ¿Por qué consideras que son un buen ajuste esos valores de “m” y “b”?, argumenta tu respuesta

g) Ahora con mismos datos que introduciste en GeoGebra, nuevamente seleccionalos e inmediatamente se desplegarán varias opciones, elije la opción “Análisis de Regresión de dos variables”. Una vez que le das aceptar y veas que se genera el diagrama de dispersión; en el menú de la parte inferior elige el modo de regresión lineal.



h) En la pantalla se despliega la gráfica en color rojo de la recta de regresión lineal y en la parte inferior también en color rojo aparece la ecuación de la misma. Escribela aquí en tu documento.



i) Compara tu resultado de la ecuación que obtuviste en el inciso “e” (línea recta supuesta) con la ecuación de la recta que obtuviste con el software en el inciso “h” (línea recta ajustada).

j) ¿Se parece o no tu resultado gráfico con el obtenido analíticamente?, argumenta tu respuesta.

k) ¿Qué signo tiene el valor de la pendiente “m” en la recta ajustada?, ¿es el mismo signo que tiene el valor de la pendiente “m” en la recta que ajustaste gráficamente.

l) En términos de población jovén, ¿qué significado tiene el signo de “m”?

m) Calcula en GeoGebra el valor de la población joven en México para el año 2020

n) ¿En qué año aproximadamente el valor de la población juvenil en México será cero?

o) ¿Cuál es un requisito indispensable para que suceda lo anteriormente cuestionado?

p) Busca en internet el concepto de *coeficiente de correlación*, contrasta la información de por lo menos dos sitios y escribe la direccción o fuente de consulta

q) Ahora en el menú de la ventana desplegada en GeoGebra selecciona visualizar estadísticas, y ahí vas a encontrar el valor del coeficiente de correlación el cual se despliega con la letra “r”

r) Una vez que ya tienes tu ajuste de la recta de regresión lineal de la población joven en México, ¿qué acciones o medidas en terminos de políticas públicas podrías sugerir que se tomen en base a los resultados obtenidos?

Anexo 2. Tablas de registros de los dados inventados por los alumnos

Kenya:

2	3	5	6	11	10	9	9	8	4	3	2	4	5	7	9	11	12	2	12

Nisa Danahe:

5	2	5	7	10	12	9	7	6	8	4	3	5	11	7	4	5	7	10	11

Daniela:

4	3	2	10	12	11	10	7	6	5	4	6	2	5	9	9	10	4	3	12

Laura Selene:

4	5	6	7	9	10	6	5	4	3	12	2	11	10	11	5	6	7	9	12

Iván:

3	4	9	11	12	2	4	5	7	8	7	9	4	3	8	11	12	5	6	4

Victoria:

3	5	7	10	2	3	5	5	8	11	12	10	9	8	7	10	4	6	7	11

Adrián:

10	2	3	5	8	11	10	12	3	5	6	7	7	9	4	7	6	11	2	12

Adriana:

3	12	2	11	3	10	4	9	11	6	7	6	7	5	7	5	4	8	8	4

Brenda:

5	2	4	8	3	11	9	4	4	7	8	7	3	6	9	5	6	4	7	8

Fernanda:

5	3	10	2	9	4	7	8	5	12	3	11	6	10	11	7	2	4	8	7

Alfonso:

2	5	3	10	2	9	3	12	3	4	5	11	4	6	8	3	9	7	4	2

Aniloreny:

7	5	2	11	10	3	5	12	9	6	7	4	5	12	8	6	4	5	6	8

Ángel:

4	7	5	9	11	12	8	9	6	8	7	3	5	4	2	11	10	3	8	12