



## La enseñanza de correlación y regresión en la formación de profesores de matemática: Un estudio exploratorio

Recepción: 09/12/2020 | Revisión: 19/02/2021 | Aceptación: 10/08/2021 | Preprint: 01/10/2021 | Publicación: 01/03/2022

 **Ana Rosa CORICA**  
CONICET - NIECyT – UNCPBA, Argentina  
acorica@exa.unicen.edu.ar  
<https://orcid.org/0000-0003-0505-0608>

 **Claudia Noemí FERRARI**  
Universidad de la Plata, Argentina  
claudianferrari@yahoo.com.ar  
<https://orcid.org/0000-0001-5063-8457>

**Resumen:** En este trabajo, con fundamento en la Teoría Antropológica de lo Didáctico, indagamos en la formación de estudiantes universitarios en Argentina futuros profesores de matemática acerca de las nociones de correlación y regresión. Reportamos resultados del análisis de la organización matemática propuesta para enseñar en la formación estadística de estos estudiantes. La descripción de la organización se realizó a partir de técnicas de revisión de documentos. Los resultados principales indican que las tareas que se proponen en la formación de profesores en matemática en torno a correlación y regresión son puntuales, rígidas y aisladas. Se requiere del estudio de tareas con datos reales y relevantes que promuevan la reflexión y el debate.

**Palabras clave:** formación de profesores; regresión; correlación; Teoría Antropológica de lo Didáctico.

### **TEACHING CORRELATION AND REGRESSION IN MATHEMATICS TEACHER TRAINING: AN EXPLORATORY STUDY**

**Abstract:** In this study, based on the Anthropological Theory of the Didactic, we investigate the training of mathematics pre-service teachers about the notions of correlation and regression in Argentina. We report the results of the analysis of the mathematical organization proposed to be taught in the statistical training of pre-service teachers. The description of the organization was made following document review techniques. The main results indicate that the tasks proposed in the training of mathematics pre-service teachers about correlation and regression are specific, rigid, and isolated. Consequently, tasks with real and relevant data are needed that promote reflection and debate.

**Keywords:** teacher training; regression; correlation; Anthropological Theory of the Didactic.

### **L'ENSENYAMENT DE LA CORRELACIÓ I REGRESSIÓ EN LA FORMACIÓ DE PROFESSORS DE MATEMÀTIQUES: UN ESTUDI EXPLORATORI**

**Resum:** En aquest treball fonamentat en la Teoria Antropològica del Didàctic, indaguem en la formació d'estudiants universitaris a l'Argentina futurs mestres de matemàtiques sobre les nocions de correlació i regressió. Reportem resultats de l'anàlisi de l'organització matemàtica proposada per a la formació estadística d'aquests estudiants. La descripció de l'organització es realitzà a partir



*de tècniques de revisió de documents. Els resultats principals indiquen que les tasques que es proposen en la formació de professors de matemàtiques vers la correlació i la regressió són puntuals, rígides i aïllades. Es necessiten, per tant, tasques amb dades reals i rellevants que promoguin la reflexió i el debat.*

**Paraules clau:** *formació de professors; regressió; correlació; Teoria Antropològica del Didàctic.*

## Introducció

Las nociones de correlación y regresión son ideas básicas en estadística, útiles en las ciencias y en diferentes campos de la actividad humana (Batanero et al., 2015). El estudio de la correlación y la regresión permite “conocer si los sucesos se relacionan y, con qué intensidad lo hacen, facilita a las personas explicar el pasado, controlar el presente y predecir el futuro” (Crocker, 1981, p. 272). Si bien, las ideas de correlación y regresión permiten la toma de decisiones en situaciones de incertidumbre, no llegan a percibirse como necesarias para la cultura estadística (Estepa et al., 2012; Gea et al., 2013). De hecho, las investigaciones didácticas sobre correlación y regresión son escasas en relación con las que refieren a la estadística descriptiva.

La relevancia y utilidad de las nociones de correlación y regresión para la formación de los ciudadanos es evidente; sin embargo, su enseñanza y aprendizaje no están exceptuados de problemas didácticos. Algunas investigaciones destacan dificultades en los estudiantes de la comprensión de nociones relacionadas con la covarianza y correlación; como ser no distinguir una distribución bidimensional de dos conjuntos de datos independientes; la concepción unidireccional de la correlación (aceptar únicamente la correlación directa); o la concepción causal (confundir correlación y causalidad) al estimar una correlación significativa (Estepa, 2008; Estepa y Batanero, 1995, 1996; Zieffler, 2006).

Con relación a la formación de profesores, las investigaciones en torno a la correlación y regresión son pocas (Gea et al., 2018). Investigaciones desarrolladas con profesores en servicio indican las dificultades de estos en la gestión del estudio de nociones de correlación y regresión (Quintas et al., 2015). La enseñanza de la correlación y regresión no es simple, pues se basa en la comprensión de diferentes conceptos como los de variable estadística y distribución bidimensional; dependencia estadística, funcional e independencia; covarianza y correlación; regresión; variable dependiente e independiente; modelo y modelo lineal; bondad de ajuste y coeficiente de determinación (Batanero et al., 2017). Las investigaciones ponen en evidencia que la enseñanza de la estadística en la escuela secundaria requiere una modificación. Siendo que la formación docente constituye la pieza fundamental para cualquier cambio del sistema educativo, el análisis de la formación estadística se torna fundamental (Stohl, 2005).

En este trabajo, con fundamento en la Teoría Antropológica de lo Didáctico (Chevallard, 1999), indagamos la formación de estudiantes para profesor de matemática de secundaria (EPM) acerca de las nociones de correlación y regresión, que estudian en instituciones terciarias en Argentina. Reportamos resultados del análisis de la organización matemática propuesta para enseñar (OMPE) en la formación estadística de EPM. Esto es esencial para, por una parte, el estudio de fenómenos relacionados con las condiciones de existencia y el tránsito de las organizaciones matemáticas entre las instituciones y, por otra, para el diseño de dispositivos didácticos que permitan abordar los fenómenos detectados.



## 1. Marco teórico

En este trabajo adoptamos como marco teórico a la Teoría Antropológica de lo Didáctico (TAD) (Chevallard, 1999). El postulado base de la TAD supone que toda actividad humana regularmente realizada se puede modelizar con la noción de praxeología (Chevallard, 1999). Esta noción es el principal constructo que se propone en la TAD para describir conocimientos y actividades en entornos institucionales. La noción de praxeología vincula el aspecto conceptual y procedimental de la actividad humana al incluir, como entidades inseparables, la *praxis* constituida por tipos de tareas y técnicas para resolverlos, y el *logos* conformado por discursos y herramientas teóricas para describir, explicar, justificar y nutrir la *praxis*. Una praxeológica u organización matemática (OM), se comprende a partir de sus elementos constitutivos. El corazón de la noción de praxeología lo constituye la tarea. Por ejemplo, calcular la mediana de un conjunto de datos es un ejemplo de tarea, mientras que calcular, es un género de tarea. Las tareas y géneros de tarea son construcciones institucionales cuya reconstrucción, en una institución determinada, es el objeto mismo de la didáctica (Chevallard, 1999). La manera de realizar una tarea se denomina técnica; esta no necesariamente es única y depende de la institución en que se realiza el estudio. El discurso racional que permite justificar y explicar a la técnica empleada al realizar la tarea se denomina tecnología. Esta tiene además como función aportar elementos para desarrollar la técnica, con la finalidad de ampliar su alcance, superar sus limitaciones y hacer posible la producción de nuevas técnicas. El último nivel de justificación de la actividad matemática lo constituye la teoría y resulta ser la tecnología de esta tecnología.

En la TAD se clasifican las OM según las características de sus componentes. Se comprende a una OM como puntual (OMP), en una determinada institución, cuando es generada por un único tipo de tarea. La integración de diversas OMP en torno a un discurso tecnológico común en una institución genera una praxeología local (OML). Mientras que una praxeología se denomina regional (OMR) en una institución si se obtiene a partir de la coordinación, articulación e integración de diversas OML, en torno a una teoría común. Las OM globales son producto de la agregación de OMR.

Fonseca (2004) establece las características que deben poseer las OM para poder integrarse en una OML, puesto que los sistemas de enseñanza deberían, al menos procurar reconstruir una OML. El cumplimiento de estas características da cuenta de cierto grado de completitud de las OM. En el proceso de estudio de una OML relativamente completa se distinguen dos partes: una relativa al proceso de construcción o reconstrucción de la propia OM determinada por los momentos didácticos (Chevallard, 1999), y otra, relativa al propio producto resultante. En particular, en lo que concierne al estudio del producto del proceso de construcción, se realiza en relación a los siguientes indicadores (Fonseca, 2004; Lucas, 2010), y que son empleados como parte de la descripción que se propone en este trabajo:

**OML1.** Integración de los tipos de tareas y existencia de tareas relativas al cuestionamiento tecnológico; esto es, tareas que hagan referencia a la interpretación, la justificación, la fiabilidad, la economía y el alcance de las técnicas, así como a la comparación entre ellas.

**OML2:** Diferentes técnicas para cada tipo de tareas y criterios para elegir entre ellas. Este indicador comporta la existencia de los elementos tecnológicos que



permiten discernir, para cada tarea concreta, cuál es la técnica más fiable y económica para llevar a cabo dicha tarea.

**OML3:** Independencia de los objetos ostensivos que sirven para representar las técnicas. La flexibilidad de las técnicas utilizadas debe permitir el uso de diferentes representaciones, pero también deben existir criterios explícitos para elegir la representación más adecuada, dependiendo de la actividad matemática en la que estas técnicas se encuentran inmersas.

**OML4:** Existencia de tareas y de técnicas "inversas". La flexibilidad de las técnicas debe también permitir trabajar tareas inversas como, por ejemplo, aquellas definidas intercambiando los datos y las incógnitas del problema o, a partir de la respuesta, analizar la situación de partida.

**OML5:** Interpretación del funcionamiento y del resultado de aplicar las técnicas. Debe existir un tipo de tarea que permita al alumno interpretar el real funcionamiento de una técnica para, a posteriori, percibir su beneficio matemático o ventaja en relación con otras técnicas.

**OML6:** Existencia de tareas matemáticas "abiertas". El estudiante ha de decidir, ante una situación matemática o extramatemática, qué datos debe utilizar y cuáles son las incógnitas.

**OML7:** Integración de los elementos tecnológicos e incidencia sobre la práctica. La tecnología y la teoría son los componentes para la construcción de técnicas nuevas, capaces de ampliar los tipos de problemas que se pueden abordar y, en consecuencia, los tipos de tareas de una organización matemática.

**OML8:** La posibilidad de perturbar la situación inicial o modificar la hipótesis del sistema para estudiar casos diferentes permite ampliar y completar el proceso de estudio.

## 2. Metodología

La investigación que se reporta es de naturaleza cualitativa, siendo el diseño propuesto un estudio de caso (Skate, 1999). Describimos la OMPE en torno a las nociones de correlación y regresión que se propone en la formación de EPM que estudian en Institutos de Formación Docente (ISFD) de la provincia de Buenos Aires, en Argentina.

La descripción de la OMPE se la realizó a partir del empleo de técnicas de revisión de documentos (Hernández et al., 2014). Estos documentos refieren a los media sugeridos para estudiar estadística por cuatro profesores que se ocupan de formar a EPM en estadística en ISFD. Un media es cualquier sistema que represente una parte del mundo natural o social destinado a un público específico, como pueden ser: un artículo de matemática, una revista, un libro, un sitio de Internet, textos editados por los mismos profesores, etc. (Chevallard, 2007). En nuestro caso, los media sugeridos para el estudio son textos editados por los profesores y el capítulo de un libro. En estos se encuentra explícito los entornos tecnológicos – teóricos, las



tareas y las técnicas que los profesores proponen estudiar en el aula, por lo que constituyen una fuente fundamental para la reconstrucción de la OMPE.

### 2.1 Características de las instituciones y profesores que participaron de la investigación

Los ISFD son los que hoy en día sustentan gran parte de la oferta de formación docente en Argentina. En la actualidad hay 1444 ISFD que dependen de la Dirección General de Cultura y Educación de la provincia en la cual desarrollan sus acciones. La mayor concentración de ISFD se ubica en la provincia de Buenos Aires que concentra a 475 instituciones. Esta provincia cuenta con 69 ISFD que dictan la carrera Profesorado en Matemática. Estas instituciones se encuentran distribuidas en 25 unidades espaciales delimitadas de acuerdo con un programa de acción, denominadas Regiones Educativas. Para la selección de la muestra de profesores que participaron de esta investigación se basó en el criterio del investigador, generándose una muestra por conveniencia (Kelmansky, 2009) atendiendo a las posibilidades físicas de relevar datos empíricos. Los ISFD que conformaron los puntos muestrales son los emplazados en la región Educativa 4, donde hay tres ISFD de gestión estatal y tres de gestión privada.

La descripción de los OMPE que se reportan en este trabajo se hace a partir de los media propuestos para el estudio por cuatro profesores. Esto es debido a que de los seis ISFD de la Región Educativa 4, en uno de ellos el profesor manifestó no enseñar estadística por la escasa carga horaria de la materia, y en otro la investigadora era la profesora. Los profesores que facilitaron los media que emplean fueron designados como P<sub>A</sub>, P<sub>B</sub>, P<sub>C</sub> y P<sub>D</sub>. Los dos primeros ejercen funciones en ISFD de gestión estatal, mientras que los otros dos, en instituciones de gestión privada. A los fines de este trabajo, se describe la OMPE que se reconstruyó a partir del media empleado por los profesores P<sub>B</sub> y P<sub>D</sub>. Estos son los únicos profesores que proponen el estudio de nociones de correlación y regresión.

### 2.2 El diseño curricular para la formación docente en matemática

La Dirección General de Cultura y Educación de la Provincia de Buenos Aires establece el DCPBA para el profesorado en matemática. En el momento del desarrollo de la investigación, los EPM se encontraban realizando sus carreras según lo establecido por Dirección General de Cultura y Educación de la Provincia de Buenos Aires (1999). En este diseño curricular se establece que la carrera de profesorado en matemática tiene una duración de 4 años y para el tercer año de la carrera se propone la materia Probabilidad y Estadística. Esta tiene una carga horaria de 96 horas anuales, distribuidas en 32 semanas. Esto indica que se destinan 3 horas semanales para el estudio de probabilidad y estadística. En la Tabla 1 se indican las nociones de estadística que se proponen estudiar, destacándose aquellos que refieren a este trabajo.

Unidad temática	Contenidos
Estadística descriptiva	Registro y presentación de datos. Histograma. Polígono de frecuencias. Medidas de centralización. Medidas de dispersión.
Nociones sobre la teoría de las muestras	Concepto de muestra. Muestreo con y sin reemplazamiento. Distribución muestral de medias. Distribución muestral de proporciones. Distribución muestral de diferencias y sumas.



Inferencia estadística	Estimación de parámetros: población y muestra. Estimación conceptual y por intervalos. Test de hipótesis. Construcción de un test. Distribución t de Student. Distribución x de Pearson.
Dependencia estadística. Regresión y correlación	Regresión lineal. Correlación lineal. Cálculo de los coeficientes de correlación y regresión lineales. Correlación y regresión múltiples. Planos de regresión y coeficientes de correlación. Coeficiente de correlación múltiple. Correlación parcial. Relaciones entre los coeficientes de correlación múltiple y parcial. Regresión múltiple no lineal.

Tabla 1. Nociones de estadística propuestas en el DCPBA.

En la Tabla 1 se destaca la cantidad de nociones que se proponen estudiar con una escasa disponibilidad temporal. Esta propuesta del diseño curricular constituye una primera restricción para estudiar estadística en la formación de EPM. El estudio de estas nociones requiere recorrer el Modelo Praxeológico de Referencia (MPR) que desarrollamos en el siguiente apartado. La formulación de un problema de investigación en didáctica de la matemática involucra una interpretación del ámbito de la actividad matemática que se encuentra implicada, y esto se sintetiza en el MPR. Este modelo es considerado como una hipótesis provisional a contrastar de manera experimental, por lo que es susceptible de ser modificado y revisado. La descripción que se hace del MPR es una red de cuestiones y respuestas donde éstas tienen estructura praxeológica.

### 3. Modelo praxeológico de referencia

La formulación del problema de investigación sobre la enseñanza de la correlación y la regresión en el marco de la TAD incluye el estudio del proceso de transposición didáctica y, por tanto, de las condiciones y limitaciones derivadas de sus diferentes elementos. En el primer paso del proceso, la pregunta es sobre cómo se conciben las nociones de correlación y regresión en instituciones académicas como ser las formadas por matemáticos. Abordar esta dimensión del problema de investigación implica la construcción de un MPR que haga explícito el punto de vista del investigador (Chevallard, 2012, García et al., 2019). Este modelo constituye una herramienta para analizar las condiciones y restricciones que pueden favorecer o dificultar el estudio de la correlación y regresión en la formación de profesores en matemática. El MPR consta de una red de praxeologías matemáticas cuya dinámica comporta ampliaciones y completaciones progresivas. Estos modelos son relativos y provisionales abiertos a cuestionamiento y revisión, resultando fértiles para la identificación de fenómenos didácticos y la formulación de problemas didácticos (Gascón, 2014).

El MPR construido se compone de un conjunto de OM que atienden tanto al análisis exploratorio de datos como al análisis inferencial. Este modelo se origina a partir de la pregunta: *Q<sub>0</sub>: ¿Cómo analizar un conjunto de datos variables?* A partir de *Q<sub>0</sub>* se deriva un conjunto de preguntas que dan origen a 19 OMs articuladas entre sí (Ferrari, 2019). En la imagen 1 se indica de manera sintética el MPR propuesto.



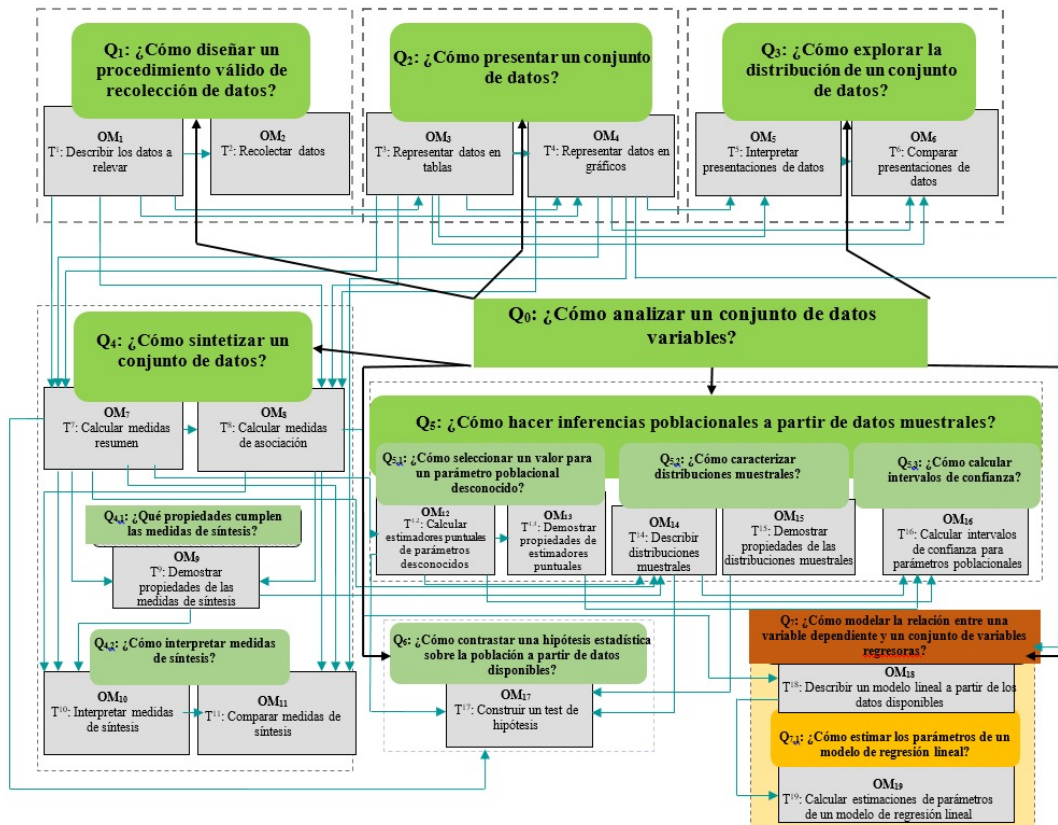


Imagen 1. Modelo Praxeológico de Referencia.

Los tipos de tareas que conforman las distintas OM que componen al MPR, se corresponden con ocho géneros de tareas que se definen a continuación, y se consideran fundamentales para el estudio de nociones estadísticas.  $G^1$ : *Recolectar*: se refiere a tareas que requieren reunir datos de variables estadísticas, registrados sobre una muestra o población;  $G^2$ : *Describir*: reúne tareas que demandan detallar las características de la información, a través de sus partes relevantes;  $G^3$ : *Calcular*: hace referencia a tareas que requieren aplicar algoritmos preestablecidos para obtener un resultado;  $G^4$ : *Representar*: congloera tareas que implican informar con tablas o gráficos el comportamiento de los datos;  $G^5$ : *Comparar*: indica tareas que se refieren a establecer relaciones de semejanza o diferencia entre dos conjuntos de datos, contrastando sus representaciones por gráficas, tablas o sus medidas resumen;  $G^6$ : *Interpretar*: agrupa las tareas que refieren a llevar a cabo el análisis de datos a partir de tablas, gráficos y/o medidas resumen, que permita extraer conclusiones, explicitando el significado de ciertas nociones, en relación a las mismas;  $G^7$ : *Demostrar*: indica tareas que requieren probar afirmaciones utilizando una secuencia finita de deducciones lógicas que se inician con axiomas o resultados demostrados; y finalmente  $G^8$ : *Construir*: se refiere a tareas que implican conducir los pasos de un proceso que permite tomar decisiones con los datos disponibles. En particular, el género de tarea  $G^6$ : *Interpretar* es transversal a todo el modelo: no es posible concebir el estudio de la estadística sin hacer interpretaciones de las nociones, dentro de la propia estadística y en el contexto de los datos que se analizan.



En lo que respecta a las nociones de interés para este trabajo, en el esquema de la Imagen 1, las OM se destacan con tonos naranjas. Este sector del MPR se origina a partir del estudio de la pregunta  $Q_7$ : *¿Cómo modelar la relación entre una variable dependiente y un conjunto de variables regresoras?* Esta pregunta se encuentra vinculada al estudio de datos de varias variables numéricas registrados sobre los mismos individuos. La pregunta  $Q_7$  genera el estudio de la  $OM_{18}$  caracterizada por el tipo de tareas  $T^{18}$ : *Describir un modelo lineal a partir de los datos disponibles*. Se intenta predecir el valor que tomará una variable continua denominada respuesta o variable dependiente, conociendo la relación que la vincula con otras variables continuas llamadas regresoras, predictoras, explicativas o independientes. El caso más simple consiste en describir una relación lineal entre dos variables continuas a partir de un conjunto de datos bivariados registrados sobre los mismos individuos. Sin embargo, esta situación rara vez acontece con datos reales; en forma más realista se puede plantear que el valor esperado de  $Y$ , cambia linealmente con  $X$ . Pero otras causas producen que los valores observados de  $Y$  varíen alrededor de su media cuando  $X$  toma un valor específico. Por ese motivo, se plantea el modelo de regresión lineal simple que incluye un término de error para registrar esas causas de variabilidad en la variable dependiente y el cumplimiento de un conjunto de supuestos. En ocasiones, especialmente en función del conocimiento teórico del problema o del diagrama de dispersión de la variable respuesta versus la variable explicativa, es necesario ajustar un polinomio de grado mayor que uno. En estos casos, la regresión lineal se denomina polinomial simple.

El entorno tecnológico asociado a  $OM_{18}$  se gesta en el hacer de la  $OM_4$  representada por el tipo de tareas  $T^4$ : *Representar datos en gráficos* y la  $OM_8$  representada por el tipo de tareas  $T^8$ : *Calcular medidas de asociación*, junto con nociones probabilísticas como esperanza condicional, distribución normal y propiedades de la esperanza y la varianza de variables aleatorias.

Una ampliación de la  $OM_{18}$  es la  $OM_{19}$ . La pregunta generatriz de esta OM resulta ser una subpregunta de  $Q_7$ ,  $Q_{7,1}$ : *¿Cómo estimar los parámetros de un modelo de regresión lineal?* Está conformada por el tipo de tareas  $T^{19}$ : *Calcular estimaciones de los parámetros de un modelo de regresión lineal*. Con la intención de predecir el valor esperado de la variable respuesta, conocidos los valores de las regresoras, se necesita estimar los parámetros del modelo. Los coeficientes del modelo se estiman a partir de la muestra aleatoria de  $n$  observaciones y serán los parámetros de la recta (regresión lineal simple) o plano (regresión múltiple con dos regresoras) que *mejor ajusta* a los datos observados. Para encontrar las estimaciones de los parámetros del modelo, postulamos que la recta o plano sea tal que minimice las distancias verticales entre los puntos observados y los puntos que están sobre la recta o sobre el plano. Los valores que hacen mínima esta distancia serán los estimadores mínimos cuadrados de los parámetros del modelo. El entorno tecnológico que justifica el hacer de la  $OM_{19}$  incluye nociones de cálculo diferencial, debido a la necesidad de minimizar funciones de dos variables, además del hacer de las tareas de la  $OM_{18}$  y las definiciones de valor predicho y residuo.

#### 4. Análisis de la organización matemática propuesta para enseñar

La descripción de la OMPE se hace a partir del análisis de los medios sugeridos para el estudio por los profesores que participaron de la investigación. Los medios que proponen los profesores son capítulos de libros o textos editados por ellos. Estos son un producto del proceso de



transposición didáctica, es decir, de la adaptación del conocimiento matemático para ser enseñado en una institución particular. En general, contienen en el siguiente orden: enunciados que corresponden al entorno tecnológico - teórico, ejemplares de tareas con indicaciones de las técnicas necesarias para su hacer y la propuesta de tareas para su estudio.

La descripción de los media propuestos por los profesores  $P_B$  y  $P_D$  se la hizo a través de las categorías que se organizaron en una tabla de análisis. Para esto, se identificó las tareas que componen a cada media, y para cada una de ellas se las analizó identificando las siguientes categorías: *género de tareas* al que refiere la tarea, *tipo de tareas* con el que se identifica la tarea, *número de tarea* según se indica en el media, *entorno tecnológico-teórico* inmediato necesario para el hacer de la tarea e *indicador de completitud matemático* al que refiere la tarea. El estudio requirió el análisis de los ejemplares de tareas explícitos en el material para interpretar las técnicas empleadas y el entorno tecnológico - teórico que justifica su hacer; y la resolución de cada tarea propuesta para los estudiantes empleando técnicas en correspondencia con el entorno tecnológico - teórico explícito en el media y atendiendo al hacer explícito en los ejemplares de tareas resueltos.

#### 4.1 Media sugerido para el estudio por el profesor $P_B$

El profesor  $P_B$  sugiere un texto para el estudio de la matemática en la escuela secundaria destinado a estudiantes de 16 años (Vázquez de Tapia et al., 1983). El capítulo propuesto para el estudio de la estadística se estructura en torno a las siguientes nociones: *población; individuo; atributos y su clasificación; serie simple y de frecuencias; frecuencias, tipos; tablas de frecuencia; gráficos estadísticos; medidas de posición: media, mediana y moda; asimetría; medidas de dispersión: amplitud; desviación; desviación media; varianza; desviación estándar; muestras; correlación y regresión lineal.*

Los ejemplares de tareas resueltos y las tareas que se proponen para que los estudiantes resuelvan pertenecen a los siguientes géneros de tareas, atendiendo a la misma definición que se estableció en el MPR: *Describir; Calcular; Representar; Comparar e Interpretar.* Esto se concluye a partir de la confección de la tabla de análisis para cada una de las tareas propuestas en el material. Por su extensión, en este trabajo nos ocuparemos de describir las características de los ejemplares de tareas resueltos en el media que dan cuenta de las técnicas propuestas para su estudio, junto al entorno tecnológico teórico asociado al hacer.

Con relación a las nociones de interés para este trabajo, una vez definidas las nociones de *ley de regresión y diagrama de dispersión*, en el texto se proponen cuatro ejemplares de tareas resueltos. Estas tareas se asocian a los tipos de tareas: *Representar datos en gráficos e Interpretar las presentaciones de datos.* Las nociones demandadas para el hacer de las mismas resultan ser: *ley de regresión, diagrama de dispersión y correlación.* El indicador de completitud asociado a estos ejemplares de tarea es  $OML_1$  siendo que las tareas en su conjunto hacen referencia a la interpretación, la justificación, la fiabilidad, la economía y el alcance de las técnicas.

En el ejemplar 1 (Imagen 2) se indica la relación entre la medida del lado de un cuadrado ( $X$ , en mm) y su perímetro ( $Y$ , en mm). Se representan los pares de puntos de la tabla y se señala, a partir del diagrama de dispersión, que existe entre las variables  $X$  e  $Y$  una perfecta correlación lineal funcional. Esta correlación es directa pues los puntos se encuentran sobre una recta de pendiente positiva.

**Ejemplo 1:**

Relación entre el lado del cuadrado y su perímetro.  
 X: lado del cuadrado medido en mm.  
 Y: perímetro del cuadrado medido en mm.

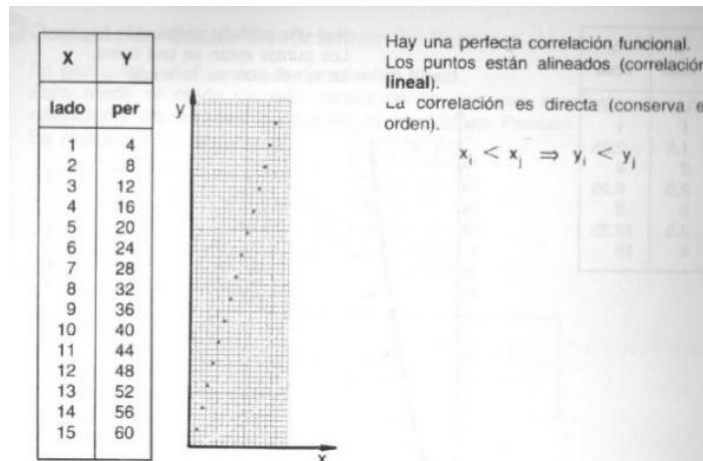


Imagen 2. Correlación lineal perfecta y directa. Fuente: Media propuesto por PB (Tapia et al., pp. 388-389).

En el ejemplar 2, (Imagen 3) se indica el diagrama de dispersión asociado a los puntos de una tabla en la que se proponen valores para la base (X, en mm) y la altura (Y, en mm) de rectángulos de perímetro constante e igual a 96 mm. Los puntos se encuentran sobre una recta de pendiente negativa lo que permite señalar que existe una correlación lineal funcional inversa.

**Ejemplo 2**

Relación entre la base y la altura de un rectángulo cuando se mantiene el perímetro constante. (Per. rect. = 96 mm).  
 X: base del rectángulo (en mm).  
 Y: altura del rectángulo (en mm).



Imagen 3. Correlación lineal perfecta e inversa. Fuente: Media propuesto por PB (Tapia et al., p. 389).

En el ejemplar 3 (Imagen 4) se indica una tabla de valores correspondientes al lado (X, en mm) de un cuadrado y el área (Y, en mm<sup>2</sup>). El diagrama de dispersión asociado muestra que los puntos se disponen sobre un arco de parábola: la correlación es funcional, aunque no lineal.

#### Ejemplo 3

Relación entre el lado del cuadrado y su área.  
 X: lado del cuadrado medido en mm.  
 Y: área del cuadrado medida en mm<sup>2</sup>.

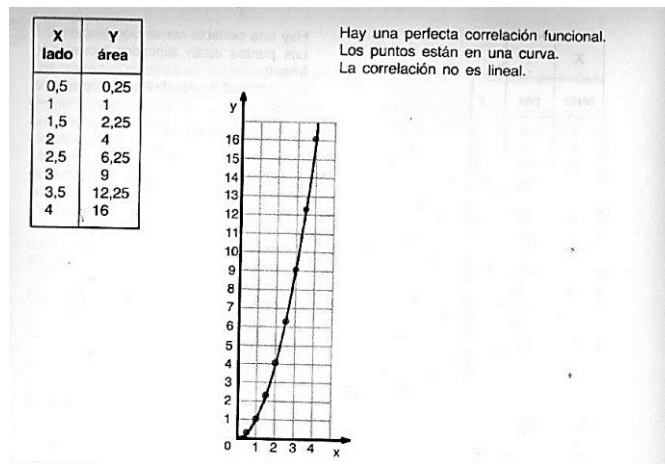


Imagen 4. Correlación funcional no lineal. Fuente: Media propuesto por PB (Tapia et al., p. 389 - 390).

El ejemplar 4 (Imagen 5) se indica una tabla de valores del número de lados de un conjunto de polígonos (X) y su correspondiente perímetro (Y). El diagrama de dispersión asociado a la tabla muestra una nube de puntos que no se encuentran ubicados sobre alguna curva particular. Aquí se indica que no hay correlación.

#### Ejemplo 4

Relación entre el perímetro y el número de lados en un conjunto de polígonos.  
 X: número de lados  
 Y: perímetro

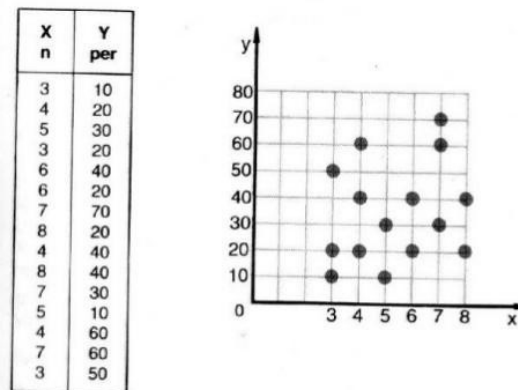


Imagen 5. Diagrama de dispersión. Fuente: Media propuesto por PB (Tapia et al., p. 390).

A partir del análisis de los cuatro ejemplares de tarea, se destaca que no se presenta una tarea que involucre una nube de puntos que pudiera ser aproximada por una curva sin que corresponda a una correlación lineal o una correlación funcional no lineal perfecta. Tampoco se exhibe la necesidad del diagrama o de la determinación del tipo de relación existente entre las variables. Podría usarse la presencia de relación lineal o cuadrática de los tres primeros ejemplares para hacer estimaciones de valores de alguna de las variables a partir de valores conocidos de la otra variable.

Una vez definido el *coeficiente de correlación* y su expresión analítica, se presentan tres ejemplares de tareas resueltas indicados en las imágenes 6, 7 y 8. Cada ejemplar exhibe datos simulados de altura y alcance (distancia entre las puntas de los dedos mayores con brazos extendidos) de un conjunto de 5 niños. Los tipos de tareas a los que refiere cada ejemplar de tarea son: *Representar datos en gráficos*, *Interpretar las presentaciones de datos*, *Calcular medidas de asociación e Interpretar medidas de síntesis*. La realización de estas tareas requiere las nociones de *ley de regresión*, *diagrama de dispersión*, *correlación* y *coeficiente de correlación*. Estos ejemplares de tarea se asocian a los indicadores de completitud, OML<sub>1</sub>, OML<sub>3</sub> y OML<sub>5</sub>, siendo que en las tareas se procura la interpretación del funcionamiento y del resultado de la aplicación de las técnicas, la justificación, la fiabilidad, la economía y el alcance de la técnica. Además, se proponen diferentes representaciones en el hacer de la tarea.

El ejemplar 1 (Imagen 6) exhibe la tabla de valores de las variables X: altura, e Y: alcance, sus medias aritméticas y aquellos valores requeridos para el cálculo del coeficiente de correlación.

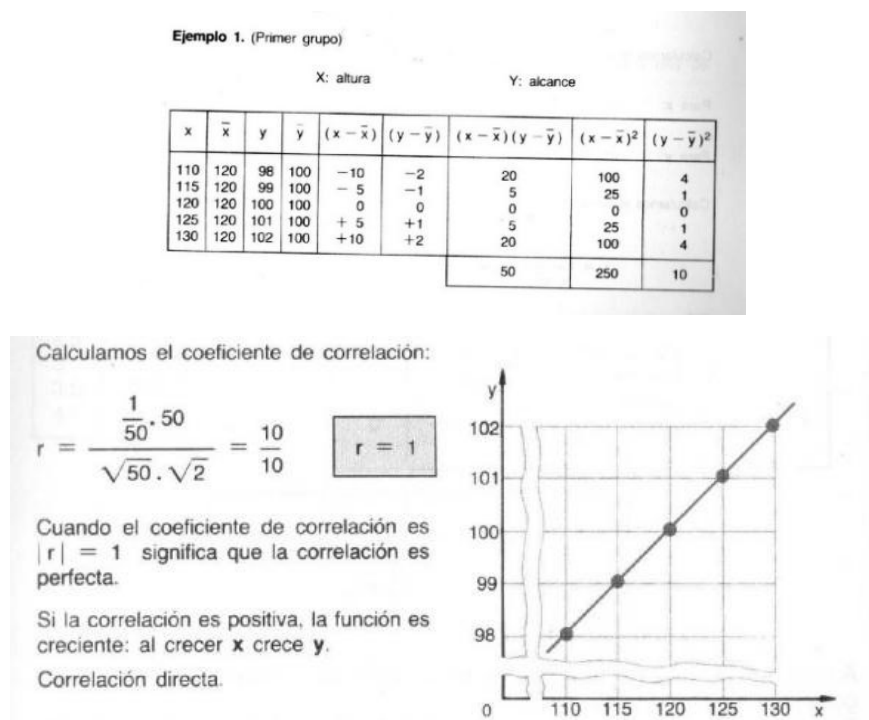


Imagen 6. Coeficiente de correlación perfecta y directa.  
 Fuente: Media propuesto por P<sub>B</sub> (Tapia et al., p. 391 -392).

El valor del coeficiente de correlación  $r = 1$  indica que la correlación es perfecta y directa; este hecho se visualiza en el diagrama de dispersión, el que muestra los puntos ubicados sobre una recta de pendiente positiva. En el ejemplar 2 (Imagen 7) se indica la tabla con la información del segundo grupo de niños. Se muestran los datos registrados y los valores que se requieren para el cálculo del coeficiente de correlación.

**Ejemplo 2. (Segundo grupo)**

x	$\bar{x}$	y	$\bar{y}$	$(x - \bar{x})$	$(y - \bar{y})$	$(x - \bar{x})(y - \bar{y})$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
110	120	102	100	-10	+2	-20	100	4
115	120	101	100	-5	+1	-5	25	1
120	120	100	100	0	0	0	0	0
125	120	99	100	+5	-1	-5	25	1
130	120	98	100	+10	-2	-20	100	4
						-50	250	10

Calculamos la desviación estándar para ambas variables.

Para x:  $\sigma_x^2 = \frac{250}{5} = 50 \Rightarrow \sigma_x = \sqrt{50}$

Para y:  $\sigma_y^2 = \frac{10}{5} = 2 \Rightarrow \sigma_y = \sqrt{2}$

Calculamos el coeficiente de correlación:

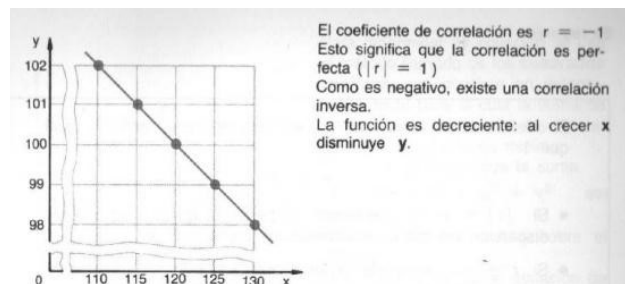
$$r = \frac{\frac{1}{5} \cdot (-50)}{\sqrt{50} \cdot \sqrt{2}} = \frac{-10}{10} = -1 \quad \boxed{r = -1}$$


Imagen 7. Coeficiente de correlación perfecta e inversa. Fuente: Media propuesto por PB (Tapia et al., pp. 392 - 393).

Dado que el valor del coeficiente de correlación  $r = -1$  se indica que la correlación es perfecta e inversa; este hecho se visualiza en el diagrama de dispersión, el que muestra los puntos ubicados sobre una recta de pendiente negativa.

El ejemplar 3 se indica en la Imagen 8. Dado que el valor del coeficiente de correlación es  $r = 0,3$ , no existe una ley de regresión entre las variables.

**Ejemplo 3. (Tercer grupo)**

x	$\bar{x}$	y	$\bar{y}$	$(x - \bar{x})$	$(y - \bar{y})$	$(x - \bar{x})(y - \bar{y})$	$(x - \bar{x})^2$	$(y - \bar{y})^2$
110	120	100	100	-10	0	0	100	0
115	120	98	100	-5	-2	+10	25	4
120	120	102	100	0	+2	0	0	4
125	120	99	100	+5	-1	-5	25	1
130	120	101	100	+10	+1	+10	100	1
						+15	250	10

La desviación estándar de las variables es la misma de los ejemplos anteriores.

$$\sigma_x = \sqrt{50} \quad \sigma_y = \sqrt{2}$$

El coeficiente de correlación es:

$$r = \frac{\frac{1}{5} \cdot 15}{\sqrt{50} \cdot \sqrt{2}} = \frac{3}{10} = 0,3 \quad \boxed{r = 0,3}$$

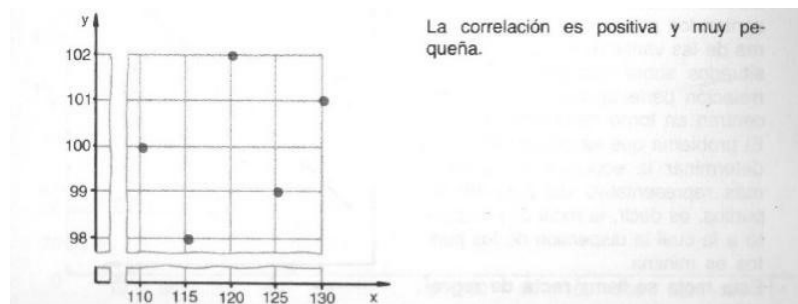


Imagen 8. Coeficiente de correlación. Fuente: Media propuesto por PB (Tapia et al., p. 393).

Los tres ejemplares se caracterizan por la aplicación de técnicas rutinarias ya sean algebraicas o gráficas, pero no existe interpretación de los coeficientes de correlación establecidos en el contexto de los datos. No se evidencia la utilidad del análisis con fines predictivos.

Luego de la noción de *recta de regresión* y de su expresión analítica, se proponen tres ejemplares de tareas resueltas. Se utilizan los datos de los tres ejemplares anteriores, es decir, los datos de altura y alcance de los tres conjuntos de cinco niños. Los tres ejemplares exhiben idénticas resoluciones. Por una parte, se indica la ecuación de la recta de regresión encontrada por simple utilización de su expresión analítica y, por otra parte, se muestra el diagrama de dispersión junto al gráfico de la recta de regresión. Estos ejemplares se indican en la Imagen 9.

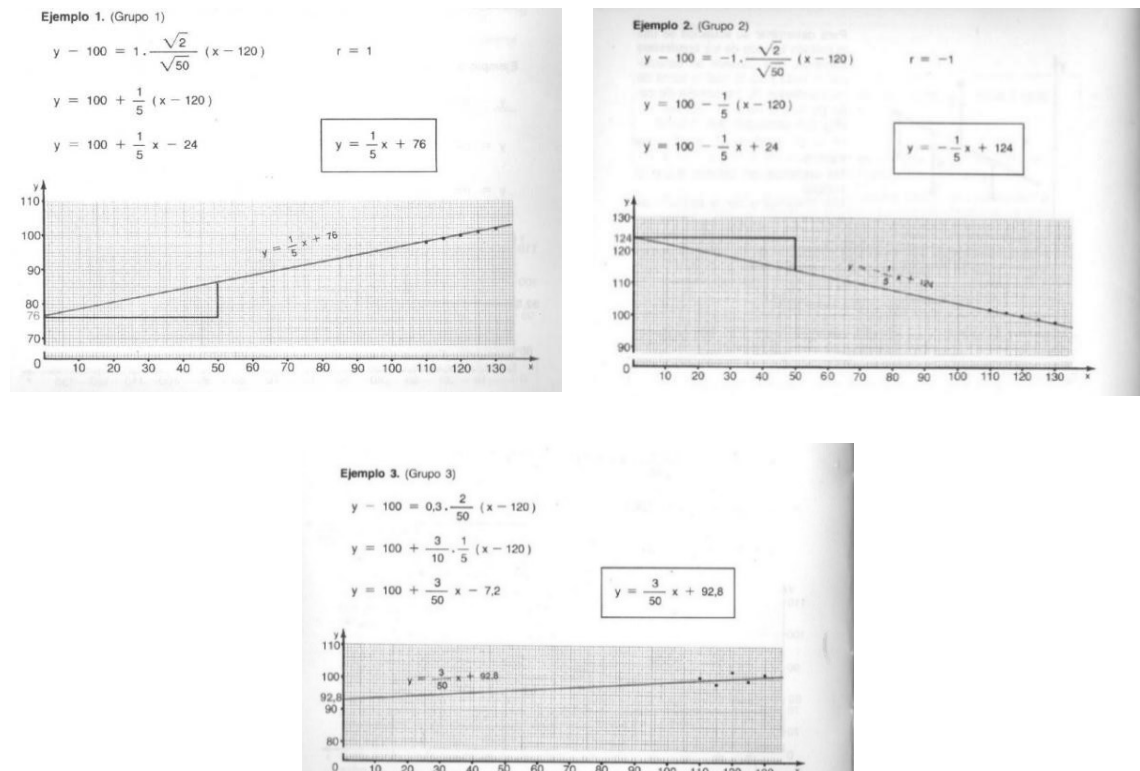


Imagen 9. Recta de regresión. Fuente: Media propuesto por PB (Tapia et al., pp. 395 - 396).



Los tres ejemplares refieren al mismo tipo de tareas *Calcular estimaciones de parámetros de un modelo de regresión lineal*. El hacer de estos ejemplares demanda de la noción de *recta de regresión*. Se destacan en los tres gráficos la ordenada al origen y la pendiente de la recta de regresión; sin embargo, no se exhibe ninguna interpretación de estos valores en el contexto de los datos. No se utiliza la recta de regresión con fines predictivos. Estos ejemplares de tarea se asocian a los indicadores de completitud, OML<sub>1</sub>, OML<sub>3</sub> y OML<sub>5</sub>, siendo que en la tarea se procura la interpretación del funcionamiento y del resultado de la aplicación de las técnicas, la justificación, la fiabilidad, la economía y el alcance de la técnica. Además, se proponen diferentes representaciones en el hacer de la tarea.

A partir de la información recogida, se evidencia la ausencia de la realización de *tareas inversas*, dado que no es posible asociar el indicador OML<sub>4</sub> a ninguna tarea de las propuestas. Tampoco se proponen tareas abiertas donde los datos e incógnitas no estén establecidos desde un principio (OML<sub>6</sub>) o el estudio de tareas que conduzcan a realizar conjeturas y definir un nuevo medio de estudio (OML<sub>8</sub>). Esta presentación es coherente con las tareas que se proponen para ser realizadas por los estudiantes. Esto se deriva del análisis de la confección de la tabla de análisis para las tareas que se proponen en el material. Por su extensión, no se presenta en este trabajo pero puede ser consultada en Ferrari (2019).

#### 4.2 Media sugerido para el estudio por el profesor P<sub>D</sub>

El profesor P<sub>D</sub> sugiere el empleo de un media que se estructura en torno a las nociones estadísticas de: *variables y su clasificación; población, muestra e individuo; distribución de frecuencias para datos agrupados y no agrupados; medidas de posición central: media aritmética, media geométrica, mediana y moda; medidas de posición no central: cuartiles, deciles y percentiles; medidas de dispersión: rango, desvío medio, varianza, desvío estándar, coeficiente de variación; gráficos estadísticos; regresión y correlación lineal*.

Los ejemplares de tareas resueltos y las tareas propuestas a los estudiantes, pertenecen a algunos de los siguientes géneros de tareas: *Describir, Calcular, Representar, Comparar e Interpretar*, como fueron definidos en el MPR.

En primer lugar, una vez presentada la noción de *curva de ajuste*, se propone el ejemplar de tarea resuelto de la Imagen 10. Este ejemplar refiere a los tipos de tareas *Representar datos en gráficos e Interpretar las presentaciones de datos*. Las nociones que demanda el hacer de esta tarea son *diagrama de dispersión y curva de ajuste*. La observación de las presentaciones da cuenta del tipo de relación entre las variables consideradas.

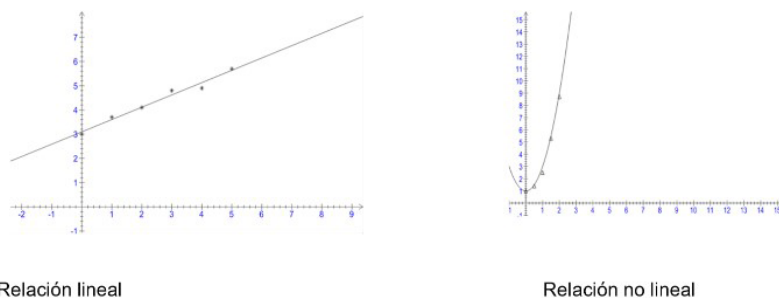


Imagen 10. Curva de ajuste. Fuente: media editado por el P<sub>D</sub> (p. 440).

Se destaca que los datos representados en ambos gráficos de la Imagen 10 no se encuentran contextualizados. Los indicadores de completitud asociados son OML<sub>5</sub> y OML<sub>7</sub>, en que se procura la interpretación del funcionamiento y del resultado de aplicar las técnicas y, además, se procura la integración de los elementos tecnológicos y su estudio en la incidencia sobre la práctica. A continuación, se presentan ecuaciones normales de la recta de regresión y se propone el ejemplo de la Imagen 11. Si bien se trata de una tarea propuesta para que realicen los estudiantes, los datos que se proporcionan en la tarea son involucrados en el hacer del ejemplar de tarea indicado en la Imagen 18.

#### EJEMPLOS

a) Sea la siguiente tabla de datos, que vinculan las variables X e Y:

X	1	3	4	6	8	9	11	14
Y	1	2	4	4	5	7	8	9

- 1) Ubicar los puntos en un sistema de coordenadas rectangulares
- 2) Ajustar libremente una recta de aproximación, y calcular su ecuación
- 3) Comparar los valores de Y de la tabla con los estimados según la ecuación obtenida. ¿cuál será el valor estimado  $Y_{est}$  para  $X = 10$ ?
- 4) Ajustar a los datos de la tabla una recta de mínimos cuadrados utilizando X como variable independiente.
- 5) Ídem utilizando Y como variable independiente.
- 6) Representar en un mismo gráfico las dos rectas de regresión obtenidas. Verificar que el centro de gravedad es la intersección de ambas.
- 7) Estimar los valores de Y para  $X = 3$ ;  $X = 6$  y  $X = 12$ .
- 8) Estimar los valores de X para  $Y = 2$ ;  $Y = 4$  y  $Y = 8$ .
- 9) Recalcular la ecuación de las rectas de regresión utilizando el *corolario*.

Imagen 11. Tarea propuesta en el material de PD. Fuente: Fuente: media editado por el PD (p. 442).

Los datos que se proponen estudiar en la tarea no se encuentran contextualizado, lo que dificulta la interpretación de los resultados solicitados en la tarea. El inciso 1) refiere al tipo de tareas *Representar datos en gráficos* y su hacer demanda la noción de *diagrama de dispersión*. La ubicación de los puntos en un sistema de ejes cartesianos da por resultado el siguiente gráfico (Imagen 12).

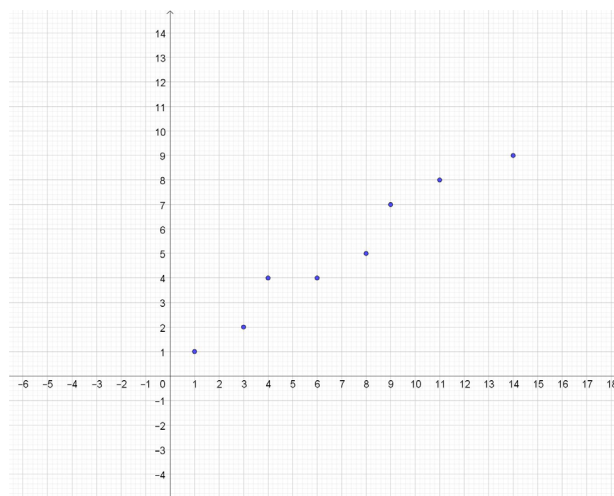


Imagen 12. Diagrama de dispersión.

El inciso 2) de la tarea de la Imagen 11 se asocia al tipo de tareas *Describir un modelo lineal a partir de los datos disponibles*. El entorno tecnológico-teórico para el hacer de esta tarea se conforma con las nociones de *curva de ajuste y regresión*. La posibilidad de que cada uno de los estudiantes proponga una curva de ajuste es notablemente enriquecedora; la justificación de la elección de la recta de ajuste dará cuenta de la necesidad de encontrar una estrategia que facilite la determinación de la mejor ecuación, es decir de la *mejor curva de ajuste*. En la Imagen 13, se presenta una de las posibles respuestas a este inciso.

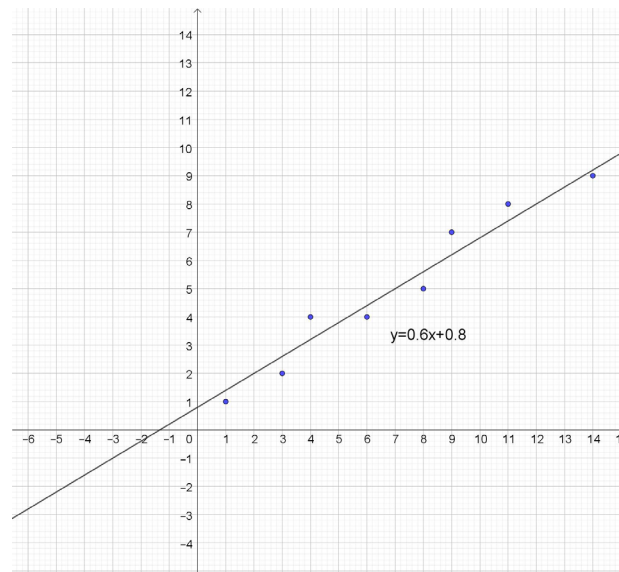


Imagen 13. Curva de ajuste.

La ecuación requerida se obtiene eligiendo dos puntos que determinen la recta que *más se acerque* a los datos y depende de la elección de cada estudiante. Así, proponiendo los puntos A y B que *a ojo* se encuentren en la nube del diagrama de dispersión, se encuentra la ecuación de la recta que se determina. Considerando, por ejemplo, los puntos A(2;2) y B(7;5), la recta de ecuación  $y=0,6x+0,8$  se considera como una recta de ajuste libre.

La realización de la tarea del inciso 3) de la Imagen 11 refiere al tipo de tareas *Calcular estimaciones de parámetros de un modelo de regresión lineal* y demanda las nociones de *regresión y error o residuo*. Aquí se demanda la determinación de los coeficientes de la recta de ajuste. En la Tabla 2 se presentan los valores observados de la variable dependiente y las correspondientes estimaciones mediante la recta propuesta en el inciso anterior. De la tabla se puede concluir que, a excepción de los valores estimados para  $x=9$  y  $x=11$ , la recta de ajuste sobrestima valores de Y.

X	1	3	4	6	8	9	11	14
$Y_{obs}$	1	2	4	4	5	7	8	8
$Y_{est}$	1,4	2,6	3,2	4,4	5,6	6,2	7,4	9,2

Tabla 2. Valores observados y valores estimados.

La estimación del valor de la variable dependiente, cuando la variable independiente asume el valor 10, se obtiene a partir de una función lineal; se destaca que  $x=10$  se encuentra dentro del rango de valores relevados para X. Por lo tanto, si  $x=10$ , entonces  $Y(10)=6,8$ .

Los incisos 4) y 5) de la Imagen 11 refieren al tipo de tareas *Calcular estimaciones de parámetros de un modelo de regresión lineal*. La realización de ambas tareas demanda las nociones de *mejor curva de ajuste y recta de mínimos cuadrado*. La aplicación de la técnica inmediatamente antes descrita conduce a la expresión  $y=0,545 + 0,636x$ , para el caso de la recta de regresión de Y sobre X y, a la expresión  $x=-0,5 + 1,5y$ , para la regresión de X sobre Y. Aquí no se solicita la interpretación de los coeficientes de las rectas de regresión, posiblemente, por la falta de contextualización de los valores registrados.

La realización de la tarea del inciso 6) de la Imagen 11 demanda la construcción de rectas, y la determinación del punto de intersección de las rectas de regresión refiere al tipo de tareas *Calcular medidas resumen* y su realización demanda la noción de *media aritmética*. Ambas rectas se representan en el gráfico de la Imagen 14.

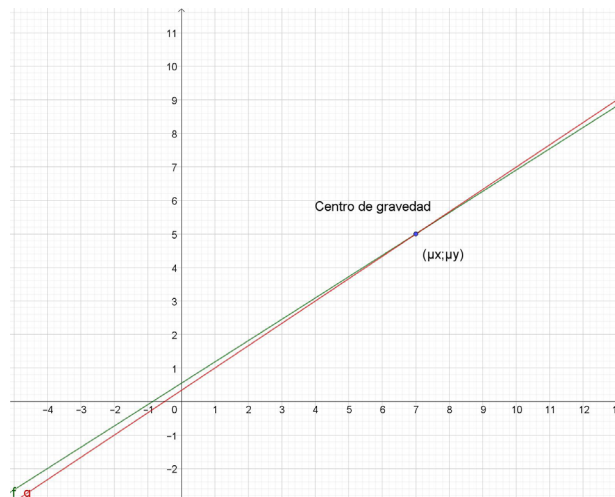


Imagen 14. Recta de regresión de y sobre x y de x sobre y.

En el gráfico se visualiza el punto de intersección de ambas rectas cuyas coordenadas son los valores medios de ambas variables. Los incisos 7) y 8) de la Imagen 11 refieren al mismo tipo de tareas: *Calcular estimaciones de parámetros de un modelo de regresión lineal*. La realización de las tareas demanda la noción de *regresión* y conduce a la confección de las siguientes tablas, pero la ausencia de un contexto impide que se interpreten los valores estimados.

Inciso 7)			Inciso 8)			
X	3	6	Y	2	4	8
Y est	2,45	4,36	X est	2,5	5,5	11,5

Tabla 3. Tabla de valores estimados.

En relación con la tarea del inciso 9) de la Imagen 11, no se encuentra en el material ningún corolario al que esta alude. En general, las tareas que involucran la Imagen 11 se encuentran asociadas a los indicadores de completitud OML<sub>5</sub> y OML<sub>7</sub>, siendo que se procura la interpretación del funcionamiento y del resultado de aplicar las técnicas y además, se procura la integración de los elementos tecnológicos y su estudio en la incidencia sobre la práctica.

A continuación, se amplía el entorno tecnológico-teórico con la incorporación de la noción de *coeficiente de correlación* y se propone el ejemplar de tarea de la Imagen 15 para que sea resuelta por los estudiantes.

#### EJEMPLOS

b) con referencia al ejemplo (a):

- 1) Calcular el error típico de Y sobre X
- 2) Calcular el coeficiente de correlación.

**Observación:** a continuación figuran tabulados en Excel, los datos necesarios para la resolución de los ejemplos a) y b).

Si bien no es imprescindible, lo reiterativo de los cálculos hace muy conveniente diseñar dicha planilla en Excel, y una vez verificada, puede utilizarse para la mayoría de los problemas de regresión y correlación.

ejemplo a y b - tablas										
X	Y	X <sup>2</sup>	XY	Y <sup>2</sup>	x = X - μ <sub>X</sub>	y = Y - μ <sub>Y</sub>	x <sup>2</sup>	xy	y <sup>2</sup>	
1	1	1	1	1	-6	-4	36	24	16	
3	2	9	6	4	-4	-3	16	12	9	
4	4	16	16	16	-3	-1	9	3	1	
6	4	36	24	16	-1	-1	1	1	1	
8	5	64	40	25	1	0	1	0	0	
9	7	81	63	49	2	2	4	4	4	
11	8	121	88	64	4	3	16	12	9	
14	9	196	126	81	7	4	49	28	16	
Σ	56	40	524	364	256			132	84	56
μ	7	5								

si  $Y = 0,545 + 0,636 X$ ; entonces

X	1	3	4	6	8	9	11	14	Σ
Y	1	2	4	4	5	7	8	9	
Y est	1,181	2,453	3,089	4,361	5,633	6,269	7,541	9,449	
Y - Y est	-0,181	-0,453	0,911	-0,361	-0,633	0,731	0,459	-0,449	
(Y - Y est) <sup>2</sup>	0,032761	0,205209	0,829921	0,130321	0,400689	0,534361	0,210681	0,201601	2,545544
Y est - μ <sub>Y</sub>	-3,819	-2,547	-1,911	-0,639	0,633	1,269	2,541	4,449	
(Y est - μ <sub>Y</sub> ) <sup>2</sup>	14,584761	6,487209	3,651921	0,408321	0,400689	1,610361	6,456681	19,793601	53,393544
Y - μ <sub>Y</sub>	-4	-3	-1	-1	0	2	3	4	
(Y - μ <sub>Y</sub> ) <sup>2</sup>	16	9	1	1	0	4	9	16	56

Imagen 15. Ejemplar de tarea del material de P<sub>D</sub>. Fuente: Fuente: Media editado por el P<sub>D</sub> (p. 444).

En esta tarea, se utilizan los datos de la Imagen 11. El inciso 1) refiere al tipo de tareas *Calcular medidas resumen* y su realización demanda del entorno tecnológico-teórico conformado por *error típico de la estimación*. La aplicación de la expresión para su cálculo, a partir de la información de la tabla, conduce al valor  $\varepsilon_{Y;X}=0,56408599$ . Se destaca que la interpretación de este valor no es requerida.

El inciso 2) se corresponde al tipo de tareas *Calcular medidas de asociación* y su realización demanda la noción de *coeficiente de correlación* o *coeficiente de Pearson*. Con la información de la tabla en Excel y la simple aplicación de la expresión de este coeficiente, se tiene  $\rho=0,9770$ . Se destaca que no se requiere interpretación del valor de este coeficiente; sin embargo, por ser un valor cercano a uno, se puede señalar la existencia de una alta correlación positiva entre las variables X e Y.



El indicador de completitud que se asocia a la tarea es OML<sub>7</sub> siendo que se procura la integración de los elementos tecnológicos y su estudio en la incidencia sobre la práctica.

Al finalizar el material, se proponen tareas para ser resueltas. Esta información se sintetizó a partir de la confección de la tabla de análisis. Por su extensión no se presenta en este trabajo, pero puede ser consultada en Ferrari (2019). Esta propuesta es coherente con los ejemplares de tareas analizados. La descripción de los ejemplares de tareas resueltos y las tareas intercaladas entre ellos permitió observar tareas que se pueden asociar con los indicadores OML<sub>3</sub>, OML<sub>5</sub>, OML<sub>6</sub> y OML<sub>7</sub>. En particular, en todo el material no se observan tareas que conduzcan a realizar conjeturas y definir un nuevo medio de estudio (OML<sub>8</sub>).

## Conclusiones

En este trabajo se analizó el DCPBA y el media propuesto por profesores quienes se encuentran a cargo del curso Probabilidad y Estadística en ISFD. El análisis realizado permite hacer inferencias acerca de la concepción epistemológica del saber. Este es organizado en teórico y práctico, como si existieran dos saberes, donde el saber práctico no tiene incidencia en la constitución del teórico. En primer lugar, se hace explícito todos los elementos tecnológicos necesarios para realizar las tareas que se proponen a continuación, luego se indican algunos ejemplos de tarea resueltos, donde se formulan las técnicas necesarias para resolver. Finalmente, se propone un conjunto de tareas de similares características, en las que se involucran las mismas técnicas que los ejemplares propuestos, pero no se proponen tareas que requieran cuestionar las técnicas empleadas, estudiando sus limitaciones y funcionalidad, para recurrir a otras técnicas.

Los principales resultados indican que las tareas que se proponen en la formación de profesores en matemática en torno a correlación y regresión son puntuales, rígidas y aisladas. El estudio que se propone se resume en un conjunto de tareas y técnicas formales, desarticuladas y débilmente interpretadas y justificadas. A partir del análisis de los ejemplares de tarea indicados en los media, se infiere que aplicar una técnica no incluye interpretar el resultado y discutir si fue correctamente utilizada.

El saber estadístico es presentado en los media sin razón de ser, dificultando la gestión de una actividad matemática funcional y, constituyendo en una restricción para la existencia de actividades matemáticas abiertas y flexibles. En particular, las nociones de correlación y regresión y su organización didáctica en los media carecen de utilidad trascendente e inherente. La ausencia de la utilidad trascendente se manifiesta en la formulación de tareas donde se proponen conjuntos de datos descontextualizados, o bien contextualizados pero luego de resolver no se interpretan los resultados en el contexto de la fuente de datos. Las tareas que involucran las nociones de regresión y correlación no corresponden a problemas científicos y/o tecnológicos; en general se aplican a datos ficticios generados exprofeso. Así también, la actualidad de los datos no parece ser importante, pues el P<sub>B</sub> propone un material para el estudio que tiene más de 30 años de su edición, resultando obsoletos los datos presentados para el estudio. En relación con la ausencia de la utilidad inherente, esta se pone de manifiesto en el estudio de tareas que no se vinculan con nociones de diferentes áreas de la matemática; solo se propone un estudio superficial e inmediato de lo que se requiere para poder estudiar las tareas que componen al media. Por lo indicado hasta aquí, se concluye en que la actividad





matemática que se propone en la formación de profesores en matemática en torno a las nociones de correlación y regresión es puntual, rígida y aislada.

En los media analizados no se encontraron tareas abiertas. La estadística, más allá de las nociones de correlación y regresión que convoca a este trabajo, es un campo propicio para la realización de tareas abiertas contrapuesta a una *enseñanza imitativa* (Burkhardt, 2006), basada en los problemas descontextualizados y convenientemente simplificados de los libros de texto (Alsina, 2007; Maaß, 2006). Se requiere del estudio de tareas estadísticas auténticas, con datos reales, relevantes en alguna situación real, sin una solución estipulada de antemano, que promuevan la reflexión y el debate de los estudiantes, donde sea necesario trabajar tanto con información matemática como no matemática, interpretando y validando los resultados en el contexto en que se sitúa el problema.

## Referencias

- Alsina, C. (2007). Si Enrique VIII tuvo 6 esposas, ¿Cuántas tuvo Enrique IV? El realismo en educación matemática y sus implicaciones docentes. *Revista Iberoamericana en investigación*, 43, 85-101. <https://doi.org/10.35362/rie430752>
- Batanero, C., Gea, M., Díaz-Levicoy, D. y Cañadas, G. (2015). Objetos matemáticos ligados a la regresión en los textos españoles de bachillerato. *Educación Matemática*, 27(2), 9-35. <https://bit.ly/3CHYQAW>
- Batanero, C., Gea, M., López-Martín, M. y Arteaga, P. (2017). Análisis de los conceptos asociados a la correlación y regresión en los textos de bachillerato, *Didacticae*, 1, 60-76. <https://doi.org/10.1344/did.2017.1.60-76>
- Burkhardt, H. (2006). Modelling in Mathematics Classrooms: reflections on past developments and the future. *Zentralblatt für Didaktik der Mathematik*, 38(2), 178-195. <https://link.springer.com/article/10.1007/BF02655888>
- Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques*, 19(2), 221-266. <https://bit.ly/3yzi6xZ>
- Chevallard, Y. (2007). Un concept en émergence: la dialectique des médias et des milieux. En G. Guedet e Y. Matheron, *Actes du Séminaire National de Didactique des Mathématiques* (pp. 344-366). ARDM et IREM de Paris 7.
- Chevallard, Y. (2012). *Théorie Anthropologique du Didactique & Ingénierie Didactique du Développement*. *Journal du séminaire TAD/IDD*. <https://bit.ly/3s6ezF6>
- Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, 90(2), 272-292. <https://psycnet.apa.org/doi/10.1037/0033-2909.90.2.272>
- Dirección General de Cultura y Educación de la Provincia de Buenos Aires. (1999). *Tomo II del Diseño curricular jurisdiccional para la formación docente de grado. Profesorado de tercer ciclo de la EGB y de la educación polimodal en matemática*. Gobierno de la provincia de Buenos Aires.
- Estepa, A. (2008). Interpretación de los diagramas de dispersión por estudiantes de Bachillerato. *Enseñanza de las Ciencias*, 26(2), 257-270. <http://funes.uniandes.edu.co/21921/>
- Estepa, A. y Batanero, C. (1995). Concepciones iniciales sobre la asociación estadística. *Enseñanza de las Ciencias*, 13(2), 155-170. <https://doi.org/10.5565/rev/ensciencias.4263>
- Estepa, A. y Batanero, M. (1996). Judgments of correlation in scatter plots: An empirical study of students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education*, 4, 25-41. <https://bit.ly/3iHd39n>



- Estepa, A., Gea, M., Cañadas, G. y Contreras, J. (2012). Algunas notas históricas sobre la correlación y regresión y su uso en el aula. *Números*, 81, 5-14.  
<http://www.ugr.es/~jmcontreras/pages/Investigacion/articulos/2012Numeros.pdf>
- Ferrari, C. (2019). *Enseñanza de la estadística en la formación de profesores: un estudio exploratorio desde la perspectiva de la Teoría Antropológica de lo Didáctico* (Tesis doctoral, Universidad Nacional del Centro de la Provincia de Buenos Aires).  
<https://www.ridaa.unicen.edu.ar/xmlui/handle/123456789/2084>
- Fonseca, C. (2004). *Discontinuidades matemáticas y didácticas entre la enseñanza secundaria y la enseñanza universitaria* (Tesis doctoral, Universidad de Vigo).  
[http://www.atd-tad.org/wp-content/uploads/2012/07/TESIS\\_en\\_PDF.pdf](http://www.atd-tad.org/wp-content/uploads/2012/07/TESIS_en_PDF.pdf)
- García, F., Barquero, B., Florensa, I. y Bosch, M. (2019). Diseño de tareas en el marco de la Teoría Antropológica de lo Didáctico. *Avances de Investigación en Educación Matemática*, 15, 75-94.  
<https://doi.org/10.35763/aiem.voi15.267>
- Gascón, J. (2014). Los modelos epistemológicos de referencia como instrumentos de emancipación de la didáctica y la historia de las matemáticas. *Educación Matemática 25 años*, 99-123.  
<https://www.redalyc.org/pdf/405/40540854006.pdf>
- Gea, M., Arteaga, P., Batanero, C. y Ortiz, J. (2018). Conocimiento tecnológico sobre la correlación y regresión: Un estudio exploratorio con futuros profesores. *Bolema*, 32(60), 134-55.  
<https://www.redalyc.org/articulo.oa?id=291265462012>
- Gea, M., Batanero, C., Fernández, J. y Gómez, E. (2013). Definiciones asociadas a la distribución de datos bidimensionales en textos españoles de Bachillerato. En J. Fernandes, M. Martinho, J. Tinoco y F. Viseu (Orgs.), *XXIV Seminario de Investigaçã em Educaçã Matemática* (pp. 127 - 140).  
<https://bit.ly/3lNSude>
- Kelmansky, D. (2009). *Estadística para todos: estrategias de pensamiento y herramientas para la solución de problemas*. INET.
- Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la Investigación*. McGraw- Hill.
- Lucas, C. (2010). *Organizaciones matemáticas locales relativamente completas* (Memoria de investigación, Diploma de Estudios Avanzados, Universidad de Vigo). <https://bit.ly/3ivHuiq>
- Maaß, K. (2006). What are modeling competencies? *Zentralblatt für Didaktik der Mathematik*, 38(2), 113-142. <https://doi.org/10.1007/BF02655885>
- Quintas, S., Ferreira, R. y Oliveira, H. (2015). O conhecimento didático de estatística de duas professoras de matemática sobre dados bivariados. *Bolema*, 29(51), 284-306.  
<https://www.periodicos.rc.biblioteca.unesp.br/index.php/bolema/article/view/8587>
- Skate, R. (1999). *Investigación con estudio de casos*. Ediciones MORATA.
- Stohl, H. (2005). Probability in teacher education and development. En G. Jones (Ed.), *Exploring probability in schools. Challenges for teaching and learning* (pp. 345-366). Springer.
- Vázquez de Tapia, N., Tapia de Bibiloni, A. y Tapia, C. (1983). *Matemática 4*. Editorial Estrada.
- Zieffler, A. (2006). *A longitudinal investigation of the development of college students' reasoning about bivariate data during an introductory statistics course*. (Tesis doctoral no publicada). Universidad de Minnesota, Min.