

Estadística aplicada a la Ingeniería: Una experiencia con alumnos universitarios

Nora Gatica

Jorge Leporati

Juan Renaudo

Graciela Echevarria

Facultad de Ingeniería y Ciencias Agropecuarias

Universidad Nacional de San Luis - Argentina

nimberti@fices.unsl.edu.ar; jlepo@fices.unsl.edu.ar

Resumen: *La regresión y la correlación así como también el control estadístico de calidad son dos técnicas estadísticas que se pueden utilizar para solucionar problemas comunes tanto en los negocios como en la empresa. Sin embargo en la enseñanza tradicional se pone mayor énfasis en el desarrollo matemático sin tener en cuenta la importancia de la aplicación de estos temas. El objetivo de este trabajo es dar a conocer una nueva metodología de enseñanza en temas de Estadística, con alumnos de Ingeniería. Se argumenta la importancia de introducir casos reales, buscados por los propios alumnos, que contribuyan a deconstruir e interpretar la clase, desde la doble perspectiva de los contenidos de estadística y de la interacción social.*

Palabras claves: *Regresión y correlación, control de calidad enseñanza estadística, alumnos universitarios, motivación.*

Applied to Engineering Statistics: An experience with university students

Abstract: *The regression and correlation as well as statistical quality control are two techniques that can be used to solve common problems in business and in the company. However, in traditional education greater emphasis is placed on the mathematical development without taking into account the importance of applying these themes. The aim*

of this paper is to present a new methodology for teaching Statistical issues with engineering students. The importance of introducing real case is argued, sought after by the students themselves, contributing to deconstruct and interpret the class, from the dual perspective of the contents of statistics and social interaction.

Keywords: *Regression and correlation, quality control, statistical training, university students, motivation.*

INTRODUCCIÓN

La estadística a pesar de contar con una axiomática satisfactoria, es quizás la única rama de la matemática donde prosiguen hoy día las discusiones sobre la interpretación de conceptos básicos, es decir si el profesor no es consciente de esta problemática, difícilmente pueda comprender algunas dificultades de sus estudiantes, quienes necesitan materializar en ejemplos concretos los conceptos y modelos matemáticos (Batanero 2001, p. 9). Desde otra perspectiva, la estadística es un área naciente de la educación matemática que tiene como elementos centrales a la teoría del constructivismo y la resolución de problemas. (Lavalle, Micheli y Rubio, 2006). En muchas situaciones de la vida real, tanto en el análisis exploratorio como confirmatorio de datos, se presentan problemas en los cuales existe una relación, asociación o dependencia estadística entre dos o más variables y se hace necesario encontrar la naturaleza de esa relación. Lo mismo ocurre con la concepción del control estadístico de calidad en donde el ingrediente básico es la utilización masiva del método científico y en concreto de la estadística en la planificación de recogida y análisis de los datos necesarios para la toma de decisiones tendientes a mejorar todos los procesos. La regresión, la correlación lineal y el control estadístico de la calidad, son contenidos de la Estadística que ya forman parte del currículo de enseñanza obligatoria en las carreras de Ingeniería. Su inclusión en el nivel universitario posibilita el tratamiento estadístico de datos bivariados como así también el análisis de calidad de los productos y o servicios. En el análisis de regresión estudiamos la relación que hay entre una variable dependiente y otra independiente, es decir estudiamos si una variables es un buen predictor de la otra variable, la medida de intensidad (si esta es numérica) de la relación existente entre una variable dependiente y otra variable independiente está determinada por lo que se denomina coeficiente de correlación. En el control de calidad, se trata, esencialmente, de minimizar la producción de unidades defectuosas reduciendo el tiempo que transcurre entre la ocurrencia y la detección de algún desajuste en el proceso de fabricación, así como la identificación de las causas del mismo a fin de evitar su repetición

FUNDAMENTACIÓN

Regresión es una palabra un tanto rara, la utilizan los biólogos, los médicos los psicólogos,... y suena como “ir hacia atrás” y realmente este es el verdadero significado del vocablo. Los primeros trabajos referidas con el estudio de regresión y se remontan

al siglo XIX, cuando Sir Francis Galton (1822-1917) imbricó sus dos grandes aficiones; el estudio de la herencia y la expresión matemática de los fenómenos vinculados a ella. Él fue el primero en trabajar con un conjunto de variables y designar a la relación entre dos variables un número para así obtener una medida tocante a su grado de relación. Sostenía la idea de que personas excepcionalmente altas solían tener hijos de estatura menor, mientras que personas muy bajas tenían hijos muy altos; este hecho fue enunciado por Galton como la regresión a la media, aplicable a las tallas de una generación respecto de las siguientes. La justificación que se da hoy día a esta situación es que los valores extremos de una distribución se deben en gran parte al azar (Lavalle, Micheli y Rubio, 2006).

En cuanto al otro tema que nos ocupamos, control de calidad, desde sus orígenes, probablemente el ser humano ha considerado de vital importancia el disponer de productos de alta calidad. Es de suponer que el cazador que disponía de mejores flechas obtenía más y mejores presas y que este hecho no debió pasar inadvertido a nuestros antepasados. La organización del trabajo en la era industrial ha añadido otros puntos de vista acerca del producto tales como costos, plazo de entrega, servicio postventa, seguridad, fiabilidad, etc. La prioridad asignada a los diversos conceptos ha ido evolucionando con el tiempo. Así, por ejemplo, en situaciones en las que la demanda de productos ha sido muy superior a la capacidad de oferta, la gestión empresarial se ha orientado hacia la producción y ha dado alta prioridad a la productividad, mientras que, cuando la demanda de ciertos productos ha sido menor que la capacidad de oferta, la gestión se ha orientado hacia el cliente y la calidad ha sido altamente prioritaria. En la actualidad pocos discuten la importancia estratégica de la calidad como factor de competitividad industrial en una situación de fuerte saturación y globalización de los mercados. Paralelamente, también ha ido evolucionando la etapa del desarrollo de un producto en la que se ha intentado asegurar su calidad.

LA EXPERIENCIA

Con la intención de que los alumnos de segundo año de la carrera de Ingeniería (en total 75 alumnos), al estudiar los temas de regresión, correlación y control de calidad, pudieran relacionar la teoría y la práctica con la realidad, y que además comprendieran la importancia de trabajar con datos reales, se les solicitó que desarrollaran estos temas con conjuntos de datos obtenidos por ellos mismos. Las fuentes de los mismos podían ser: datos observados por ellos, obtenidos de Internet o libros de textos o que realizaran el trabajo con datos extraídos en la producción de fábricas de la zona. Como la ciudad donde está ubicada la Facultad de Ingeniería es zona de promoción industrial, los alumnos, en general tienen relación con ellas por lo que no se les presentaría mayores inconvenientes, para poder obtener los datos.

El trabajo que debían presentar, consistía en una carátula, introducción y explicación teórica del tema, como y de que fuente extrajeron los datos, la aplicación a los datos de la técnica usando el paquete estadístico *statgraphics* y las conclusiones pertinentes. El mismo debía ser elaborado individualmente o en grupo con no más de dos personas por grupo. Dicho trabajo debía estar escrito en Word y grabado en un soporte informático

como un CD o disquete. A medida que los alumnos iban trabajando, las dudas o problemas que se les presentaban eran consultadas por ellos mismos al docente responsable mediante correo electrónico o en forma personalizada. De esta manera el trabajo era presentado y el alumno promocionaba la materia.

Entre tantos trabajos recibidos mencionamos los siguientes, extrayendo de ellos como obtuvieron los datos y las conclusiones a las que arribaron. No es nuestra intención exponer el desarrollo de las técnicas estadísticas desarrolladas por los alumnos, sino mostrar como obtuvieron los datos y los ejemplos que utilizaron en el trabajo.

Según las fuentes utilizadas, dividimos los trabajos en distintos grupos:

Grupo 1. Los datos fueron obtenidos de los Proyectos de Investigación de la Facultad. Los integrantes de los proyectos proporcionaron a los alumnos los elementos para que ellos realizaran el trabajo

Ejemplo 1 Tema: Regresión y correlación (alumnas: Jenifer y Mariana)

Este grupo de alumnas presentó un trabajo realizado en la Facultad de Ingeniería y Ciencias Agropecuarias de la Universidad Nacional de San Luis. Un grupo de investigadores pertenecientes al proyecto “*Estudio de reacciones catalíticas heterogéneas y de productos naturales*” estudia la actividad de catalizadores para la combustión de material particulado (carbón eliminado con los gases efluentes producido por la combustión interna de los motores diesel), a fin de lograr un dispositivo que elimine el dióxido de carbono despedidos por los caños de escape. Para realizar las medidas de actividad y proponer un modelo cinético que ajuste los datos experimentales se utiliza un reactor termogravimétrico, en dicho reactor se programa una velocidad de calentamiento lineal. El encargado del equipo les brindó la ayuda necesaria para la obtención de los datos de la termobalanza y les solicitó que verificaran si al variar la velocidad de calentamiento se mantenía la linealidad.

Partiendo de lo anterior, las alumnas plantearon como hipótesis que trabajando con bajas velocidades de calentamiento, la reacción es menos exotérmica, es decir que hay menor liberación de calor.

En una balanza termogravimétrica se carga una determinada masa, compuesta por carbón y catalizador, en una cápsula. Se programa para que aumente cada 1 minuto 10 y una cierta velocidad de calentamiento.

Se ejecuta el programa y comienza el calentamiento. De este modo, la balanza toma cada 1 segundo el peso de la cápsula, en función de la temperatura y el tiempo. Cuando se llega a una cierta temperatura, la masa empieza a consumirse y cuando lo hace por completo, es decir, hay un desplazamiento de la reacción hacia la derecha, entonces se detiene la ejecución del programa ya que no hay más materia prima.

La balanza brinda la lectura de los datos a partir de un tiempo cero, en este trabajo se tomaron en cuenta los datos en la primera experiencia a partir de los 60 minutos de haber comenzado la ejecución del programa y en la segunda experiencia a partir de los 15 minutos, debido a que observaron durante las experiencias que desde esos tiempos hay una mayor variación de los valores, que en períodos anteriores no se observaron, a fin de elaborar una conclusión más específica sobre la actividad catalítica.

Una de las fotos que acompañaron este trabajo fueron:

Balanza
termogravimétrica
con las compuertas
abiertas, donde se
colocan las muestras
de carbón.



Conclusión a la que arribaron: *Recomendamos a los técnicos del equipo utilizar para las mismas masas cargadas en la balanza termogravimétrica trabajar con velocidades de calentamiento bajas para mantener la linealidad de temperatura vs tiempo, ya que con velocidades altas la reacción de combustión es más exotérmica y no se conserva dicha linealidad.*

Grupo 2: Datos provenientes de empresas dedicadas a la explotación minera de recursos naturales tales como piedra, metales, áridos, etc.

Ejemplo 2. Tema: Regresión y correlación (alumnos: Jorge y Alberto)

Un grupo de alumnos presentó un trabajo cuyos datos corresponden a la evolución del comportamiento lineal entre las variables profundidad Vs rendimiento en una cantera ubicada en Nogolí – San Luis (Argentina). El análisis se llevó a cabo bajo la hipótesis que a medida que se avanza en profundidad, en excavaciones sobre los cúmulos de sedimentos de cantos rodados, se encuentra cada vez menos baches y más suelo arenoso, por consiguiente disminuye el rendimiento de la piedra triturada (adjuntan una fotografía que evidencia éste supuesto al observarse los estratos rocosos bien diferenciados.) Cabe aclarar que el rendimiento, medido en porcentaje, se toma con respecto a la cantidad en gramos de piedra útil (mayores a 6mm) sobre el peso en gramos de la muestra (una palada) por cien. Los datos fueron provistos por el sector de laboratorio de la Empresa Rovella Carranza S.A., de los registros de granulometría de la Cantera Nogolí.

Conclusión a la que arribaron: *Como dijimos en un comienzo, el modelo de regresión lineal es útil para determinar valores de una variable a partir de valores que toma la variable de la cual depende; remitiéndonos al análisis hecho podemos concluir que*

éste nos permite predecir el rendimiento de la roca útil mientras avanzamos en profundidad, siempre que éstos valores se encuentren dentro del rango que responde a dicho comportamiento, entonces podemos decir que en éste caso nos serviría hasta 20 o 30 metros de profundidad, hondura máxima de excavación que se ha realizado en ésta cantera, de modo que no podemos predecir que sucede al exceder ésta profundidad ya que no existen muestras extraídas. De todas formas, teniendo en cuenta la evidencia del modelo, a la empresa no le serviría, puesto que seguir excavando cada vez mas implica tener cada vez menos rendimiento. Se presentan una de las fotos obtenida por ellos durante la investigación:



Grupo 3. Datos obtenidos experimentalmente por los mismos alumnos

Ejemplo 3. Tema: Regresión y correlación (alumnas: Daniela y Marisa)

Este grupo presentó un trabajo cuyos datos fueron obtenidos experimentalmente de un proceso de deshidratado de manzana, la cual se encontraba pelada, limpia y cortada en mitades que fue colocada en un horno eléctrico específico. En este análisis se quiere observar la relación del peso (gr) de la manzana con respecto al tiempo (min) de deshidratado, para ello se utilizó un diagrama de dispersión para determinar luego el modelo de regresión adecuado.

Conclusión a la que arribaron: *El primer análisis que realizamos fue en base a un gráfico de dispersión en el cual observamos que la relación entre las dos variables era similar a una línea recta. Debido a esto, nos referimos a la misma como una*

*correlación lineal simple, a partir del software statgraphics. Una vez que ingresamos los datos a dicho programa, rescatamos la información útil para el análisis (tabla y gráficos). De la tabla pudimos concluir: Dado que el p-valor en la tabla es inferior a 0.01, existe relación estadísticamente significativa entre las variables para un nivel de confianza del 99%. El estadístico R-cuadrado indica que el modelo explica un 98,6087% de la variabilidad en Peso. El error estándar de la estimación muestra la desviación típica de los residuos que es 0,0453904. Este valor puede usarse para construir límites de la predicción para las nuevas observaciones. Como la correlación mide el grado de asociación entre dos variables y el coeficiente de correlación (-0.993019) nos dio en este caso cercano a menos uno, entonces se puede concluir que hay una correlación perfectamente negativa entre el peso y el tiempo de deshidratado, es decir indica una relación relativamente fuerte entre las variables. Como se puede observar, el modelo se ajusta a la realidad planteada a partir de estos datos. Como la grafica residual no presenta estructura sino que los valores están por debajo y por encima de cero en forma aleatoria, entonces aseguramos que los residuos se distribuyen en forma normal con media cero y la misma variabilidad. El modelo ajustados es: $\text{Peso} = 1,93857 - 0,000952916 * \text{Tiempo}$.*

Grupo 4: Datos obtenidos por aquellos alumnos que tuvieron acceso y colaboración en fábricas del medio.

Ejemplo 4. Tema: Control de calidad (alumnos Carlos y Gabriel)

Para realizar el análisis de Control de calidad se necesita elegir el sujeto o producto y su respectiva característica. En este caso el producto elegido fueron bobinas de “Film Stretch de Polietileno” utilizadas para empaquetamiento industrial, y la característica el coeficiente de fricción estático y dinámico. Se eligió esta característica debido a que es muy requerida por los clientes. Es muy útil a la hora del transporte, sobre todo cuando debe recorrer grandes distancias pues la misma puede caerse y dañarse provocando pérdidas tanto para la fabrica como para el cliente.

- Producto: Bobinas de “Film Stretch de Polietileno”.
- Característica: Coeficiente estático y dinámico.

La experiencia fue realizada en la fabrica ALTAPLASTICA S.A. donde los alumnos pudieron acceder a todo el equipamiento necesario, las instalaciones y máquinas de prueba del laboratorio. De esta forma pudieron apreciar y realizar la prueba de coeficiente de rozamiento estático y dinámico (c.o.f).

Las muestras de las bobinas fueron obtenidas de forma aleatoria con cierta diferencia horaria. Luego de tener todos los datos realizaron el análisis de control de calidad haciendo uso de la herramienta STATGRAPHICS.

Conclusión a la que arribaron: *Como podemos observar los datos se encuentran dentro de los parámetros estimados sin variaciones abruptas o que se aproximen demasiado a alguno de los límites superiores e inferiores, por lo tanto concluimos que el proceso se encuentra bajo control, y la producción puede continuar sin mayores percances (ver imágenes página siguiente).*



Imágenes de la maquina utilizada para las pruebas y recoleccion de datos: Plataforma de prueba donde se colocan las muestras de “Film Stretch de Polietileno”, se colocan dos retazos de la muestra y se hacen deslizar una sobre la otra, y el dispositivo obtiene los resultados buscados. Imágenes del proceso de obtención de datos en el laboratorio.

Ejemplo 5. Tema: Control de calidad. Alumno Alejandro

Para realizar el análisis de control de calidad se debe elegir el sujeto o producto y sus respectivas características a analizar.

En este caso el producto es un pañal para niños, las variables a analizar son:

- 1) La distribución del AGM (peso total y seccionales)
- 2) El Patch (es el componente que ayuda a la distribución uniforme del orín en el pañal)

Estas variables afectan directamente la absorción del pañal y se pueden medir por lo tanto realizaron el análisis de los mismos.

La experiencia fue realizada en la fábrica Procter & Gamble S.A.. El alumno pudo acceder a las instalaciones y maquinas de prueba en la línea de producción, donde se realizó la sección del pañal y se obtuvieron las muestras que son parte vital de este trabajo.

Las pruebas se realizaron mediante un proceso manual en el cual la fábrica le provee al operario todos los intervalos en los cuales puede oscilar la calidad del pañal y sus componentes.

Las muestras del pañal fueron obtenidas de forma aleatoria, dicho muestreo se realizó cada dos horas.

Luego de obtener todos los datos necesarios, realizó un análisis de control de calidad haciendo uso de la herramienta STATGRAPHICS.

Conclusión a la que arribó: Como podemos observar en el caso de la distribución del AGM, el proceso se encuentra bajo control dentro de los parámetros de cada zona. Cabe destacar que las Zonas 1 y 2 son las Zonas frontales del pañal en las que el AGM se tiene que encontrar en mayor proporción para evitar el traspaso del orín a la ropa. Este aditivo de forma cristalina similar al azúcar cumple una función vital en el pañal la cual es transformarse en un gel gelatinoso al entrar en contacto con el orín.

Para el caso del Patch los datos se encuentran dentro de los parámetros sin variaciones que indiquen la detención del proceso de producción. El Patch, es uno de los elementos mas comunes que se pueden encontrar en un pañal, si bien su función es distribuir el orín uniformemente en todo el pañal, no es ni más ni menos que 100% cartón molido.

Con este análisis de variables podemos decir que el proceso se encuentra bajo control y la producción puede continuar sin problemas (ver imágenes página siguiente).

Grupo 5: Son aquellos trabajos donde los datos se obtuvieron recurriendo a Internet. Por ejemplo:

Ejemplo 6. Tema: regresión y correlación. Alumno Javier

Un alumno de Ingeniería Química presentó el siguiente trabajo: Se consideró un inóculo (pequeña cantidad de un producto que contiene bacterias) de *Lactobacillus casei* de 60 ml de volumen y se lo dejó incubar por 3 horas a temperatura, humedad y acidez normales. Finalizado este período, se realizó el recuento de células cada 1 hora, por un lapso de 10 horas, sin aplicar agitación, y relacionó tiempo de incubación (hs.) con crecimiento de células (cel/ml). Los datos fueron extraídos de la siguiente página web <http://sites.google.com/site/enalcahe/microbiologia-predictiva>

Grupo 6: En este grupo se encuentran los trabajos cuyos datos son extraídos de libros de texto.

Ejemplo 7. Tema: regresión y correlación. Alumna: Laura

Esta alumna presentó un trabajo donde los datos resultan de 10 mediciones extraídos de una empresa que fabrica productos químicos. Se estudió la influencia del tiempo de extracción en la eficiencia de una operación de extracción. Los datos fueron extraídos del libro *Probabilidad y Estadística para Ingenieros* de Miller y Freud . Quinta Edición. Ejercicio 11.1 página 343.



Estiramiento en
la zona de cuchillas.



Seccionamiento.



Zona de
pesado de pañales

CONCLUSIONES

Los alumnos se sintieron entusiasmados y motivados por esta nueva metodología (de acuerdo a sus respuestas a un cuestionario que desarrollaron posteriormente donde se les preguntó sobre el trabajo realizado).

Con ayuda de bibliografía, apoyo docente, soporte informático y software estadístico, los estudiantes fueron capaces de elegir y crear sus propios datos extraídos de distintas fuentes. Mediante esta metodología se sintieron atraídos a introducirse en la investigación dando lugar a que ellos con esmero, tesón y constancia, presentaran trabajos como los citados anteriormente

Cabe destacar que la mayoría de los trabajos se encuentran en el grupo 3 (60 %) donde los datos fueron obtenidos experimentalmente por los mismos alumnos. Hay dos trabajos en el grupo 5 y solo uno en el grupo 6, por lo que concluimos que los estudiantes prefieren encontrar los datos experimentalmente y no extraerlos de libros de texto o Internet.

REFERENCIAS

- Batanero, C. (2001): *Didáctica de la Estadística*. Departamento de Didáctica de la Matemática de la Universidad de Granada. Editorial de la Facultad de Ciencias. España.
- Lavalle, A., Micheli, E. y Rubio, N. (2006): Análisis didáctico de regresión y correlación para enseñanza media. *Revista Latinoamericana de Matemática Educativa (Relime)*. 9, 383–406.
- Chillemi, M. y Morales, E. (2002): Enseñanza de correlación y regresión lineal simple. Una experiencia en carreras de Ingeniería. *Acta Latinoamericana de Matemática Educativa*. 15, 477- 483. México.
- Barletta, B. y Cadoche, S. (2002): *Los malentendidos en las conclusiones estadísticas*. Acta Latinoamericana de Matemática Educativa. 15, 484- 488. México.
- Franzini, D., Muñoz, N. y Sánchez, R. (2002): *Análisis de los errores metodológicos en trabajos escolares de Estadística*. Acta Latinoamericana de Matemática Educativa. 15, 459-464. México.
- Vallecillos, A. (2002): *Análisis de aprendizajes en inferencia estadística a través de proyectos de investigación*. Acta Latinoamericana de Matemática Educativa. 15, 453-458. México.