

**Entrevista con el matemático Víctor Yohai (UBA-CONICET) (realizada el 23/10/14, un día antes del dictado del Seminario del IMAL)**

*Tema: "Estimadores robustos altamente eficientes para muestras finitas"*

"Para poder explicar el objetivo de mi conferencia debo comenzar explicando qué son los métodos robustos. En general, los métodos clásicos en Estadística se han desarrollado bajo una hipótesis muy fuerte: el modelo que se está usando para explicar un conjunto de datos es satisfecho exactamente por todos sus elementos. Sin embargo, los procedimientos obtenidos bajo estos supuestos son muy sensibles a esta hipótesis. Basta que un pequeño porcentaje de observaciones deje de satisfacer el modelo supuesto (aun en una única observación) para que los resultados de aplicar los procedimientos clásicos sean afectados de manera muy significativa y que el consiguiente análisis de los resultados obtenidos conduzca a conclusiones erróneas. Un método estadístico es robusto si, por el contrario, tiene la propiedad de ser muy poco afectado por unas pocas observaciones que no satisfacen el modelo".

"Uno de los modelos más utilizados en Estadística es el modelo de regresión lineal. Este modelo supone una relación lineal entre una variable  $Y$  y un conjunto de variables explicativas  $X_1, \dots, X_k$  de las cuales depende la variable  $Y$  y que permiten su predicción. En este modelo, la predicción no es exacta ya que está afectada de un error, generalmente pequeño. Los métodos clásicos suponen, además, que esos errores tienen una distribución normal y bajo esa suposición derivan que la forma óptima para ajustar el modelo es el método de mínimos cuadrados. Este método hace un ajuste minimizando la suma de los cuadrados de los errores. Sin embargo, basta que una observación tenga un comportamiento atípico desviándose del modelo para que el ajuste obtenido por este procedimiento esté muy alejado del modelo correcto y, por lo tanto, inutilizable para realizar predicciones confiables".

"La conferencia tratará sobre los diferentes métodos que se han propuesto para ajustar un modelo de regresión lineal que simultáneamente tienen las siguientes dos propiedades: (a) son robustos, es decir no son afectados por un porcentaje pequeño de observaciones atípicas, y (b) en el caso que no haya observaciones atípicas y todos los errores son normales se comporta en forma muy similar al método de mínimos cuadrados.

**Trasladados a la vida práctica, ¿en qué podemos aplicar eventualmente estos estimadores robustos?**

Un ejemplo de modelo de regresión lineal donde se podrían utilizar los métodos robustos podría ser el caso de que la variable  $Y$  fuese el rendimiento por hectárea de una semilla de trigo y las variables que expliquen ese rendimiento fuesen, por ejemplo,  $X_1$  la cantidad de fertilizante utilizado por hectárea,  $X_2$  la cantidad de riego utilizado por hectárea y  $X_3$  el tipo de semilla o variedad de trigo utilizado. En caso que se utilizase el método de mínimos cuadrados, bastaría que se cometiera un error en los datos correspondientes a un pequeño porcentaje de parcelas cultivadas para que el ajuste obtenido sea altamente afectado. Esto no ocurriría si se utilizase un método robusto.

*Entrevistó: Lic. Enrique A. Rabe (ÁCS).*

*Foto: genileza Diego Sklar*