Departamento Académico de Estadística
Río Hondo # 1, Col. Progreso Tizapán,
México, D. F., C.P. 01080

Seminario Aleatorio

Sesión 283

Bayesian nonparametric hypothesis testing for the two-sample problem

Luis Gutiérrez
Escuela de Salud Pública, Facultad de Medicina, U. de Chile

Jorge González
Departamento de Estadística, Pontificia Universidad Católica de Chile

Andrés F. Barrientos
Duke University, Department of Statistical Science

Resumen

The two-sample problem is defined as the comparison between two populations on the basis of two independent samples, one from each population. In the field of biostatistics, it is of interest to identify whether a certain feature acts in the same way in both populations. Formally, let $Y_{1,1}, \dots, Y_{1,n_1}$ and $Y_{2,1}, \dots, Y_{2,n_2}$ be two samples measuring such feature, with G_1 and G_2 being the corresponding distribution functions. This problem has been traditionally addressed by evaluating if $G_1(Y_1) = G_2(Y_1 + \theta)$ so that any comparison to be made between the two populations then depends on the parameter θ. However, the above approach can be embedded in a more general setting by replacing θ by $\Delta(Y_1)$, where $\Delta(\cdot)$ is referred to as the shift function. This extended approach can be used not only to test the hypothesis that $G_1 = G_2$ but also to make inferences about the set $\{y: \Delta(y) \neq 0\}$, which allows, for instance, to identify the members of the population for which a given treatment is beneficial. We propose to develop a novel and fully Bayesian nonparametric (BNP) model to make inferences on shift functions. Specifically, we develop a procedure to test the hypothesis that $G_1 = G_2$ and to make inferences about $\{y: G_1(y) \neq G_2(y)\}$, we also extend the procedures by considering confounders and predictors. We develop eficient Markov chain Monte Carlo schemes for exploring the corresponding posterior distributions. Comparisons based on simulations between the proposal and the parametric and nonparametric alternatives commonly used in Biostatistics are provided. Finally, the proposal is illustrated with simulated and real-life datasets.

Viernes 20 de noviembre 2015, 13:00 hrs.
Aula B-4, Plantel Río Hondo

El Seminario Aleatorio está destinado tanto a profesores como a estudiantes, por lo que el Departamento de Estadística agradece a los profesores que colaboren invitando a sus alumnos a estas sesiones.

En la red*:* http://estadistica.itam.mx/es/51/contenido/seminario-aleatorio-de-estadistica