

# Teorías de información – 543727

## Miniproyecto N° 1

Jorge E. Pezoa  
[jpezoa@udec.cl](mailto:jpezoa@udec.cl)

28 de septiembre de 2018

Este miniproyecto se debe trabajar en grupos de máximo dos personas y debe entregarse a más tardar el viernes **12/10/2018 a las 23:59 hrs.** vía email a [jpezoa@udec.cl](mailto:jpezoa@udec.cl).

### Teoría

Derive los estimadores de máxima verosimilitud de los parámetros de las siguientes distribuciones: Rayleigh de parámetro  $\sigma$  y Uniforme Continua, de parámetros  $a$  y  $b$ . Además, resuelva el problema 1.1 del libro.

### Regresión lineal

Utilizando el archivo `SkillCraft1.Dataset.csv`, se pide diseñar regresores lineales usando máxima verosimilitud y Bayesianos. Para el caso Bayesiano, calcular además las distribuciones *a posteriori* de los parámetros y de la variable estimada. Además, diseñar un otro tipo de regresor lineal o no lineal que le parezca usando las herramientas de Matlab, R, Weka, python u otro software de su agrado. Mayor información sobre el conjunto de datos puede encontrarse acá: <https://archive.ics.uci.edu/ml/datasets/SkillCraft1+Master+Table+Dataset>. Para leer el conjunto de datos se recomienda el comando `csvread`.

Se recomienda los siguientes toolboxes de Matlab: *Statistics* <http://www.mathworks.com/help/stats/index.html>, *Curve Fitting* (<http://www.mathworks.com/help/toolbox/curvefit/> ver comando `cftool`). En particular, los siguientes comandos son útiles para visualización de datos: `gscatter`, `scatter`, `scatter3`, `gplotmatrix`. Considerar también los siguientes comandos para las regresiones: `lasso`, `ridge`, `fitlm`, `stepwiselm`, `fitglm`, `stepwiseglm`, `mnrfit`, `predict` y la clase `LinearModel`. Finalmente, considere el ejemplo de regresión Bayesiana <http://www.mathworks.com/products/statistics/examples.html?file=/products/demos/shipping/stats/bayesdemo.html>

La presentación de sus resultados de esta parte debe hacerse en formato póster. El póster debe contener al menos las siguientes secciones: (i) resumen, (ii) muy breve descripción del problema, (iii) solución(es) al problema, donde se debe incluir la selección de variables, tipo de regresión(es), tipo de función usada, etc. y se puede incluir, si es que sirve, diagramas de las variables, clústers, etc., (iv) resultados, y (v) conclusiones. Los resultados de los problemas se entregan en hojas adicionales.

Plantillas de pósters en L<sup>A</sup>T<sub>E</sub>X pueden encontrarse acá <http://cursos.die.udec.cl/~jpezoa/cursos/PosterLaTeX.rar> y acá [http://cursos.die.udec.cl/~jpezoa/cursos/poster\\_NUC.rar](http://cursos.die.udec.cl/~jpezoa/cursos/poster_NUC.rar). Plantillas de pósters en Microsoft PowerPoint pueden encontrarse acá <http://www.writing.engr.psu.edu/posters.html> y acá [http://www.posterpresentations.com/html/free\\_poster\\_templates.html](http://www.posterpresentations.com/html/free_poster_templates.html).