

Consideren la clase del martes 12 de marzo de 2019.

**1. Complementen con reflexiones propias las observaciones mencionadas en las secciones 2.3.4 y 3.1 del laboratorio.**

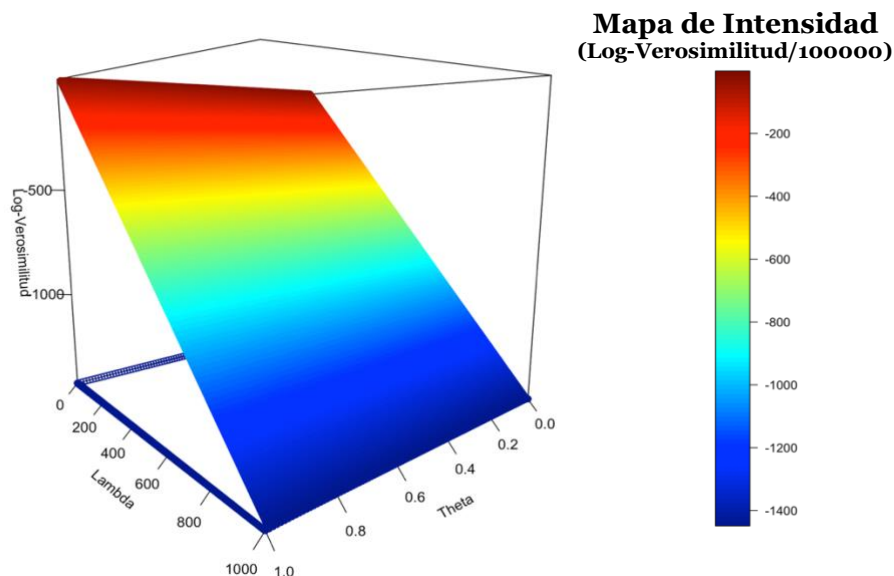
• **2.3.4:**

- Podemos observar que el tercer cuartil muestral aumenta año con año por lo que podríamos concluir que cada vez se acumulan montos mayores al 75% de nuestra distribución, es decir, cada año la mayor parte de los montos de siniestros se acumulan cada vez más en cifras más elevadas. De la misma manera, analizando el número de suscriptores durante cada año se mantiene a la alza mientras que el número de siniestros fue descendiendo a lo largo de los años.

• **3.1:**

- Complementando la idea del inciso anterior, podemos observar que los datos presentan distintas tendencias a lo largo de 2005, 2006 y 2007 por lo que usar los mismos parámetros de  $\theta$  y  $\lambda$  para los tres años no hace mucho sentido pues como ya había mencionado anteriormente, los datos no se comportan de manera homogénea entre los años. Lo adecuado sería estimar un  $\theta$  y un  $\lambda$  para cada año con la finalidad de captar las distintas tendencias que muestran los datos año con año.

**2. Respecto a la sección 3.3, grafiquen la verosimilitud (o logverosimilitud) para  $(\theta, \lambda)$  con base en los datos de 2005-2007. Grafiquen también las curvas de nivel de la verosimilitud e identifiquen clases de equivalencia estadística para los parámetros.**



- El código de esta gráfica se encuentra en el archivo “act11302\_156165\_t05.Rmd” a partir de la línea 288 hasta la 300.

**3. Respecto a la sección 3.4, comenten acerca de los supuestos que propician que el EMV (estimador máximo verosímil) de  $(\theta, \lambda)$  conjunto pueda calcularse por separado. Comenten qué condiciones harían que el cálculo no pueda hacerse por separado.**

- Para poder estimar los parámetros  $\theta$  y  $\lambda$  por separado necesitamos que las variables aleatorias correspondientes a los montos de los siniestros ( $X_j$ ) y al número de siniestros ( $t_j$ ) sean independientes e idénticamente distribuidas puesto que el hecho de que las variables no estén relacionadas nos permite resumir la información correspondiente a la frecuencia de siniestros en el parámetro  $\theta$  y la relacionada al monto de los siniestros en  $\lambda$ . Si las variables fueran dependientes, no podríamos realizar el cálculo de manera separada pues la información que contendría cada parámetro dependería de las dos variables aleatorias.