1. (a) Sea $\mathbf{X} = (X_1, \ldots, X_n)$ una muestra aleatoria con distribucion P_θ. Supongamos una distribución a priori T para θ, donde $T \sim \tau$. Sea $\delta_\Lambda(\mathbf{X})$ el estimador Bayes para θ. Supongamos que $\delta_\Lambda(\mathbf{X})$ es un estimador insesgado para θ. Probar que

$$E((\delta_\Lambda(\mathbf{X}) - T)^2) = 0$$

.

   (b) Mostrar que $\bar{X}$ no es un estimador de Bayes de θ para ninguna distribución a priori Λ cuando $X|_{T=\theta} \sim N(\theta, 1)$ y cuando se usa la pérdida cuadrática.

2. Consideremos una m.a. $X_1, \ldots, X_n$ tal que $X_i|_{=\theta} \sim \mathcal{P}(\theta)$ y $\sim \Gamma(r, \lambda)$, con $r, \lambda > 0$.

   (a) Encontrar el estimador Bayes δ_Λ y calcular $r(\delta_\Lambda, \Lambda)$ el riesgo de Bayes.

   (b) Mostrar que δ_Λ puede escribirse como un promedio pesado entre $\bar{X}$ y $\dfrac{r}{\lambda}$. Interpretar.

3. (Empirical Bayes)

   (a) Vamos a trabajar con el dataset `Batting` del paquete `Lahman`. Solo vamos a considerar las columnas `playerID`, `H` (Hits) y `AB` (Intentos). Para empezar, obtener el dataset que contenga el numero total de hits y de intentos por jugador.

   (b) Si tuviera que elegir el mejor jugador (en terminos del que hittea mas frecuentemente), ¿elegiria a los que tienen promedio 1? ¿Por que?

   (c) Considerando unicamente a los jugadores con mas de 1000 intentos, estimar la distribucion de los promedios de hitteo de los jugadores usando una distribución Beta. Llamemos Λ a esta distribucion. Sugerencia: función `fitdistr` del paquete `MASS`, o usar método de momentos.

   (d) Supongamos que la probabilidad REAL de hitteo del jugador i es p_i. Fijar la distribucón estimada del item anterior como la priori de los valores p_i. (Notar que en esta versión de Bayes, la distribución a priori no genera un único parámetro a estimar, sino que genera muchos parámetros, uno para cada jugador). Para cada jugador i, obtener el estimador Bayes del parámetro p_i.

   (e) Rankear a los jugadores con las probabilidades estimadas p_i (usando los estimadores Bayes del item anterior).

   (f) Hacer un grafico comparando dos estimaciones de p_i: primero, usando la estimación frecuentista (proporción de hitteos) y segundo, usando la estimación de Empirical Bayes del item anterior.