Chapter 10

Estimadores de Bayes

10.1 Enfoque Bayesiano del problema de la estimación puntual

Consideremos nuevamente un problema estadístico de estimación paramétrico. Se observa un vector $\mathbf{X} = (X_1, ..., X_n)$ (que puede ser por ejemplo aunque no necesariamente una muestra aleatoria de cierta destribución) con densidad discreta o continua en la familia $f(\mathbf{x}, \boldsymbol{\theta})$, con $\boldsymbol{\theta} = (\theta_1, ..., \theta_k) \in \Theta \subset I\!R^k$.

El enfoque llamadado frecuentista que hemos estudiado no supone ningún conocimiento previo de $\boldsymbol{\theta}$. El enfoque bayesiano por lo contrario supone que se tiene alguna información previa sobre $\boldsymbol{\theta}$. Esta información está expresada por medio de una distribución sobre $\boldsymbol{\theta}$, denominada distribución a priori. Aqui supondremos que esta distribución a priori tiene una densidad $\gamma(\boldsymbol{\theta})$. Está distribución a priori puede tener distintas interpretaciones según el problema. Se pueden dar las siguientes alternatvas

- La distribución a priori está basadas en experiencias previas similares.
- La distribución a priori expresa una creencia subjetiva.

El hecho de que el enfoque bayesiano considere una distribución de probabilidades sobre $\boldsymbol{\theta}$, supone tratar a $\boldsymbol{\theta}$ como una variable aleatoria, y por lo tanto a esta variable la denominaremos Θ para distinguirla del valor que toma $\boldsymbol{\theta}$. Esta notación puede llevar a confusión dado que también llamamos Θ al conjunto de valores de $\boldsymbol{\theta}$. Sin embargo por el contexto quedará claro el significado de este símbolo en cada caso.

Dada que consideramos ahora el valor del parámetro como el valor de una variable aleatoria, la interpretación de la familia de densidades $f(\mathbf{x}, \boldsymbol{\theta})$ en el enfoque bayesiano también cambia. En el enfoque bayesiano $f(\mathbf{x}, \boldsymbol{\theta})$ se interpreta como la distribución condicional de la muestra $\mathbf{X}$ dado que la variable Θ toma el valor $\boldsymbol{\theta}$.

Una vez observada la muestra $\mathbf{X}$ se puede calcular la distribución condicional de de Θ dada $\mathbf{X}$. Esta distribución se denomina distribución *a posteriori* y está dada por

$$f(\boldsymbol{\theta}|\mathbf{x}) = \frac{f(\mathbf{x}, \boldsymbol{\theta})\gamma(\boldsymbol{\theta})}{\int \ldots \int f(\mathbf{x}, \mathbf{t})\gamma(\mathbf{t})d\mathbf{t}}. \tag{10.1}$$

En efecto el numerador de (10.1) corresponde a la densidad conjunta de $\mathbf{X}$ y Θ, y el denominador a la densidad marginal de $\mathbf{X}$.

Si la distribución de $\boldsymbol{\theta}$ fuese discreta, habría que reemplazar las integrales del denominador por las correspondientes sumatorias. En lo sucesivo supondremos que es tanto las distribuciones de $\mathbf{x}$ y the $\boldsymbol{\theta}$ son continuas, pero el tratamiento en el caso discreto es similar.

Una de las ventajas del enfoque bayesiano es que se pueden definir en forma natural estimadores óptimos, sin necesidad de restricciones poco naturales como la de estimadores insesgados a la que debimos recurrir en el enfoque frecuentista. Para ver esto supongamos que queremos estimar $\lambda = q(\boldsymbol{\theta})$ y consideremos una función de pérdida $l(\lambda, \lambda')$ que indica el costo de estimar λ' en vez de λ. Supongamos que se tiene un estimador $\hat{\lambda} = \delta(\mathbf{x})$. Luego la pérdida será una variable aleatoria $l(q(\Theta), \delta(\mathbf{X}))$, y la pérdida esperada que llamaremos *riesgo de Bayes* está dada por

$$r(\delta, \gamma) = E(l(q(\Theta), \delta(\mathbf{X}))), \tag{10.2}$$

donde aqui la esperanza se toma con respecto a la distribución conjunta de $\mathbf{X}$ y Θ. Por lo tanto dado la distribución priori γ, un estimador óptimo será aquel que minimice $r(\delta, \gamma)$. Este estimador lo llamaremos estimdor de Bayes corrspondiente a la distribución a priori γ y será representado por δ_γ.

Llamemos

$$R(\delta, \boldsymbol{\theta}) = E(l(q(\Theta), \delta(\mathbf{X}))|\Theta = \boldsymbol{\theta}). \tag{10.3}$$

Luego $R(\delta, \boldsymbol{\theta})$ es la función de riesgo de la teoría frecuentista, y estará dada por

$$R(\delta, \boldsymbol{\theta}) = E_{\boldsymbol{\theta}}(l(q(\boldsymbol{\theta}), \delta(\mathbf{X}))) = \int (l(q(\boldsymbol{\theta}), \delta(\mathbf{x}))f(\mathbf{x}, \boldsymbol{\theta})d\mathbf{x}. \tag{10.4}$$

También se tendrá

$$r(\delta,\gamma) = E_\gamma(E(l(q(\Theta),\delta(\mathbf{X}))|\Theta)) = \int ... \int R(\delta,\gamma)\gamma(\boldsymbol{\theta})d\boldsymbol{\theta}. \tag{10.5}$$

Consideremos como caso particular la función de pérdida cuadrática, es decir

$$l(\lambda,\delta) = (\lambda-\delta)^2.$$

En este caso el estimador de Bayes sera la funcion $\delta_\gamma(X)$ minimizando el error cuadrático medio

$$E((\delta(\mathbf{X}) - q(\Theta))^2)$$

y por lo tanto de acuerdo a la teoría de esperanza condicional, éste será único y dado por

$$\delta_\gamma(\mathbf{x}) = E(q(\Theta)|\mathbf{X} = \mathbf{x}) = \int ... \int q(\boldsymbol{\theta})\gamma(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta},$$

es decir será la esperanza condicional de $q(\Theta)$ con respecto a la distrbución a posteriori de Θ.

Ejemplo 1. Sea $\mathbf{X} = (X_1, ..., X_n)$ una muestra independiente de una distribución Bi$(\theta, 1)$, y supongamos que la distribución a priori de θ sea una distribución beta(a, b), es decir con una densidad

$$\gamma(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}I_{[0,1]}(\theta). \tag{10.6}$$

Es conocido que esta distribución tiene la siguiente esperanza y varianza

$$E(\theta) = \frac{a}{a+b}, \tag{10.7}$$

$$var(\theta) = \frac{ab}{(a+b)^2(a+b+1)} = \frac{E(\theta)\,(1-E(\theta)}{a+b+1}. \tag{10.8}$$

Luego si se conoce la media y la varianza de la distribución a priori de Θ, se pueden determinar a y b. La fórmula (10.8) muestra que para un dado valor de la esperanza, la varianza depende de $a+b$, tendiendo a 0 cuando $a+b \to \infty$.

La distribución de la muestra $X_1, X_, ...X_n$ dado el valor de θ está dada por

$$f(x_1, ..x_n, \theta) = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}. \tag{10.9}$$

Luego usando (10.1) se tiene que la distribución a posteriori de θ está dada por

$$f(\theta|x_1,...,x_n) = \frac{\theta^{\sum_{i=1}^n x_i+a-1}(1-\theta)^{n-\sum_{i=1}^n x_i+b-1}}{\int_0^1 t^{\sum_{i=1}^n x_i+a-1}(1-t)^{n-\sum_{i=1}^n x_i+b-1}dt}. \tag{10.10}$$

Podemos observar que esta densidad defiere solamente en un factor que depende de $x_1,...,x_n$ de la correspondiente a una distribución beta$(a+\sum_{i=1}^n x_i, n-\sum_{i=1}^n +b)$. Como se trata de una densidad condicional respecto a $x_1,..,x_n$, podemos tratar a este factor como una constante. Dado que dos densidades que difieren en una constante son iguales, podemos concluir que distribución a posteriori de θ es beta$(a+\sum_{i=1}^n x_i, n-\sum_{i=1}^n +b)$.

Supongamo que la función de pérdida es cuadrática. Luego el estimador de Bayes,que denominaremos $\delta_{a,b}$, está dado por $E(\Theta|\mathbf{X})$, y de acuerdo a (10.7) tendremos que

$$\delta_{a,b} = \frac{T+a}{a+b+n} = \frac{n}{n+a+b}\,\frac{T}{n} + \frac{a+b}{a+b+n}\,\frac{a}{a+b}, \tag{10.11}$$

donde $T=\sum_{i=1}^n X_i$. Por lo tanto el estimador de Bayes se comporta como un promedio ponderado de dos estimadores: el IMVU $\delta_1 = T/n$ que no usa la información de la distribución a priori y $\delta_2 = a/(a+b)$ que corresponde a la esperanza de la distribución a priori y que se usaría si no se hubiese observado la muestra. También vemos que el peso asignado a δ_2 tiende a 0 cuando el tamaño de la muestra n aumenta.

De acuerdo a (10.11), el estimador de Bayes correspondiente a una distribución a priori beta (a,b) puede interpretarse como el estimador frecuentista correspondiente a una muestra de tamaño $n+a+b$ con $\sum_{i=1}^n X_i + a$ éxitos.

Observación 1. En el ejemplo anterior hemos partido de una distribución a priori beta(a,b), y hemos obtenido que la distribución a posteriori también está en la misma familia, ya que es beta$(a+\sum_{i=1}^n x_i, n-\sum_{i=1}^n x_i+b)$. Se dice entonces que la familia de distribuciones beta es la conjugada de la familia de distribuciones binomial.

Ejemplo 2. Sea $\mathbf{X} = (X_1,...,X_n)$ una muestra independiente de una distribución $\mathrm{N}(\theta,\sigma^2)$, con σ^2 conocido, y supongamos que la distribución a priori de θ sea $\mathrm{N}(\mu,\tau^2)$.

Luego la densidad de la muestra $\mathbf{X} = (X_1,...,X_n)$ dado θ esta dada por

$$f(\mathbf{x},\theta) = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta)^2\right), \tag{10.12}$$

donde $\exp(x) = e^x$. La densidad de la distribución a priori está dada por

$$\gamma(\theta) = \frac{1}{(2\pi)^{1/2}\tau}\exp\left(\frac{-(\theta-\mu)^2}{2\tau^2}\right) \tag{10.13}$$

Luego multiplicando (10.12) y (10.13), desarrollando los cuadrados y haciendo algún manipuleo algrebaico se obtiene que distribución conjunta de $\mathbf{X}$ y Θ está dada por

$$f_{\mathbf{X},\Theta}(\mathbf{x},\theta) = C(\mathbf{x},\sigma^2,\mu,\tau^2)\exp\left(-\frac{1}{\theta^2}\left(\frac{n}{\sigma^2}+\frac{1}{\tau^2}+\theta(\frac{n\bar{x}}{\sigma^2}+\frac{\mu}{\tau^2}\right)\right),$$

donde C no depende de θ. Completando cuadrados y luego de un manupuleo algebraico adicional se obtiene

$$f_{\mathbf{X},\Theta}(\mathbf{x},\theta) = C_1(\mathbf{x},\sigma^2,\mu,\tau^2)\exp\left(-\frac{1}{2D}\left(\theta-D\left(\frac{n\bar{x}}{\sigma^2}+\frac{u}{\tau^2}\right)\right)^2\right), \tag{10.14}$$

donde

$$D = \frac{1}{(n/\sigma^2)+(1/\tau^2)}. \tag{10.15}$$

Completando cuadrados en el exponenente de (10.14) se obtiene

$$f_{\mathbf{X},\Theta}(\mathbf{x},\theta) = C_2(\mathbf{x},\sigma^2,\mu,\tau^2)\exp\left(-\frac{1}{2D}(\theta-D(\left(\frac{\bar{x}}{\sigma^2}+\frac{\mu}{\tau^2}\right)\right)^2 \tag{10.16}$$

Finalmente usando (1) obtenemos

$$f(\theta|\mathbf{x}) = C_3(\mathbf{x},\sigma^2,\tau^2,\mu)\exp\left(-\frac{1}{2D}(\theta-D(\left(\frac{\bar{x}}{\sigma^2}+\frac{\mu}{\tau^2}\right)\right)^2 \tag{10.17}$$

Luego esta densidad excepto una función que depende solo de $\mathbf{x}$ corresponde a una distribución

$$\mathrm{N}\left(D\left(\frac{n\bar{x}}{\sigma^2}+\frac{\mu}{\tau^2}\right),D\right). \tag{10.18}$$

Como se trata de una distribución condicional $\mathbf{X}=\mathbf{x}$, podemos considerar a C_3 como constante. Luego la distribución a posteriori de Θ debe ser entonces dada por (10.18).

Supongamos nuevamente que consideramos una función de pérdida cuadrática. Luego, el estimador de Bayes estará dado por la esperanza condicional de Θ dado $\mathbf{X}$, y por lo tanto de acuerdo a (10.18) y (10.15) está dado por

$$\delta(\mathbf{X}) = D\left(\frac{n\bar{X}}{\sigma^2} + \frac{\mu}{\tau^2}\right) = w\bar{X} + (1-w)\mu, \tag{10.19}$$

donde

$$w = \frac{n/\sigma^2}{(n/\sigma^2) + (1/\tau^2)}.$$

Por lo tanto el estimador de Bayes es un promedio ponderado del estimador IMVU de la teoría frecuentista $\bar{X}$ y la media de la distribución a priori μ. Los pesos son inversamente proporcionales a las variancias σ^2/n y τ^2 de ambos estimadores. A medida que el tamaño de la muestra n crece, el peso del estimador basado en la información a priori tiende a 0. Es decir a medida que el tamaño de la muestra crece, la información a priori tiene menos relevancia para determinar el estimador de Bayes.

Observación 2. En este ejemplo partimos de una distribución a priori en la familia $\mathrm{N}(\mu, \tau^2)$, y obtenemos que la distribución a posteriori está dada por (10.18), y por lo tanto está en la misma familia normal. Luego la familia de distribuciónes conjugadas de la familia normal con varianza conocida es la la familia normal.

10.2 Utilización de métodos bayesiano para resolver problemas frecuentistas

En esta sección vamos a mostrar como los resultados de la teoría bayesiana pueden ser útiles, aunque no se comparta ese punto de vista. Es decir veremos que los resultados de esta teoría se pueden usar para resolver problemas que surgen de la teoría frecuentista.

Consideremos una muestra $\mathbf{X} = (X_1, ..., , X_n)$ con distribución conjunta $f(\mathbf{x}, \boldsymbol{\theta})$ donde el vector de parámetros $\boldsymbol{\theta} \in \Theta$. Supongamos que queremos estimar $\lambda = q(\boldsymbol{\theta})$ y que tenemos una función de pérdida $l(\lambda, \lambda')$. En el enfoque frecuentista un estimador $\delta(\mathbf{X})$ de λ queda caracterizado por su función de riesgo

$$R(\delta, \boldsymbol{\theta}) = E_{\boldsymbol{\theta}}(l(q(\theta), \delta(\mathbf{X})) = \int ... \int \delta(\mathbf{x}) f(\mathbf{x}, \boldsymbol{\theta}) d\mathbf{x}. \tag{10.20}$$

Como θ es desconocido, lo ideal sería encontrar un estimador $\delta^*(\mathbf{X})$ tal que dado cualquier eotro estimador $\delta(\mathbf{x}))$ se tuviese

$$R(\delta^*, \boldsymbol{\theta}) \leq R(\delta, \boldsymbol{\theta}) \quad \forall \boldsymbol{\theta} \in \Theta.$$

Como ya hemos visto al comienzo del curso estos estimadores no existen excepto en casos muy poco interesantes.

Una alternativa es comparar los estimadores por el máximo riesgo. Dado un estimador $\delta(\mathbf{X})$ de λ su máximo riesgo se define por

$$MR(\delta) = \sup_{\boldsymbol{\theta} \in \Theta} R(\delta, \boldsymbol{\theta}). \tag{10.21}$$

El criterio de comparar los estimadores por su máximo riesgo es pesimista, ya que actua como si el parámetro fuese a tomar el valor más desfavorable para el estimador. Un estimador óptimo de acuerdo a este criterio es un estimador δ^* tal que dado cualquier otro estimador δ se tiene

$$MR(\delta^*) \leq MR(\delta). \tag{10.22}$$

Definición. Un estimador satisfaciendo (10.22) se denomina *minimax.*

Vamos ver como la teoría bayesiana nos ayuda a encontrar estimadores minimax. Para eso consideremos una distribución a priori $\gamma(\theta)$. El correspondiente estimador de Bayes δ_γ satisfará que dado cualquier otro estimador δ se tiene

$$r(\delta_\gamma, \gamma) \leq r(\delta, \gamma). \tag{10.23}$$

Luego, de acuerdo a (10.5) se tendrá entonces que para cualquier estimador δ

$$\int R(\boldsymbol{\theta}, \delta_\gamma)\gamma(\boldsymbol{\theta})d\boldsymbol{\theta} \leq \int R(\boldsymbol{\theta}, \delta)\gamma(\boldsymbol{\theta})d\boldsymbol{\theta}. \tag{10.24}$$

El siguiente Teorema nos permite usar la teoría bayesiana para encontrar estimadores minimax.

Theorem 1. *Supongamos que se tiene una distribución a priori γ^* tal que el estimador de Bayes δ_{γ^*} es única y tal que $R(\delta_\gamma, \boldsymbol{\theta})$ es constante en θ. Luego δ_{γ^*} es el único estimador minimax*

Demostración. Consideremos un estimador $\delta \neq \delta_{\gamma^*}$. Vamos a mostrar que

$$MR(\delta) > MR(\delta_\gamma). \tag{10.25}$$

Supongamos por reducción al absurdo que (10.25) no se cumpla. Por hipóthesis existe C tal que $R(\delta, \boldsymbol{\theta}) = C$. Luego, como $MR(\delta_{\gamma^*}) = C$, si (10.25) no se cumple se tendrá

$$R(\delta, \boldsymbol{\theta}) \leq C \ \ \forall \boldsymbol{\theta} \in \Theta,$$

y por lo tanto

$$R(\delta, \boldsymbol{\theta}) \leq R(\delta_\gamma, \boldsymbol{\theta}) \ \ \forall \boldsymbol{\theta} \in \Theta. \tag{10.26}$$

Integrando ambos miembros de (10.26) se tiene

$$\int ... \int R(\delta, \boldsymbol{\theta}) \gamma^*(\boldsymbol{\theta}) d\boldsymbol{\theta} \leq \int ... \int R(\delta_{\gamma^*}, \boldsymbol{\theta}) \gamma^*(\boldsymbol{\theta}) d\boldsymbol{\theta}, \tag{10.27}$$

y de acuerdo a (10.5) se tiene

$$r(\delta, \gamma^*) \leq r(\delta_{\gamma^*}, \gamma^*)$$

Como δ_{γ^*} es el estimador de Bayes correspondiente a γ^*, esto implica

$$r(\delta, \gamma^*) = r(\delta_{\gamma^*}, \gamma^*)$$

y por lo tanto δ también es estimador de Bayes correspondiente a γ^* contradiciendo la hipótesis de que era único.

Ejemplo 3. Consideremos el Ejemplo 1 de estimacion bayesiana para la familia binomial usando distribuciones a priori en la familia beta(a, b) y como fúnción de pérdida la función $l(\theta, \theta') = (\theta - \theta')^2$. Luego hemos visto que el único estimador de Bayes esta dado por

$$\delta_{a,b} = \frac{T + a}{n + a + b},$$

con $T = \sum_{i=1}^n X_i$.

Si encontramos a y b tales que $R(\delta_{a,b}, \theta)$ es constante, ese estimador será minimax. Como $E_\theta(T) = n\theta$ y $var(T) = n\theta(1 - \theta)$ se tiene

$$E_\theta(\delta_{a,b}) = \frac{n\theta + a}{n + a + b}, \tag{10.28}$$

y

$$var_\theta(\delta_{a,b}) = \frac{n\theta(1 - \theta)}{(n + a + b)^2}, \tag{10.29}$$

Luego usando (10.28) y (10.29) se tiene

$$\begin{aligned} R(\delta_{a,b}, \boldsymbol{\theta}) &= E((\delta_{a,b} - \boldsymbol{\theta})^2) \\ &= var_\theta(\delta_{a,b}) + (\theta - E_\theta(\delta_{a,b}))^2 \\ &= \frac{n\theta(1-\theta)}{(n+a+b)^2} + \left(\theta - \frac{n\theta + a}{n+a+b}\right)^2 \\ &= \frac{n\theta(1-\theta) + (a+b)^2\theta^2 - 2a(a+b)\theta + a^2}{(n+a+b)^2} \\ &= \frac{(-n + (a+b)^2)\theta^2 + (n - 2a(a+b))\theta + a^2}{(n+a+b)^2}. \qquad (10.30) \end{aligned}$$

Para que (10.30) sea constante en $\boldsymbol{\theta}$, los coeficientes en $\boldsymbol{\theta}$ y $\boldsymbol{\theta}^2$ del numerador deben ser 0. Por lo tanto se debe cumplir

$$-n + (a+b)^2 = 0, \qquad n - 2a(a+b) = 0$$

La solución a este sistema de ecuaciones es $a = b = \sqrt{n}/2$, y por lo tanto el estimador de Bayes correspondiente, que sera minimax, estará dado por

$$\delta_{\text{mmx}} = \frac{T + (\sqrt{n}/2)}{n + \sqrt{n}}. \qquad (10.31)$$

La correspondiente función de riesgo está dado por

$$R(\delta_{\text{mmx}}, \theta) = \frac{n/4}{(n+\sqrt{n})^2} - \frac{1}{4(\sqrt{n}+1)^2}.$$

Vamos ahora a dar un segundo Teorema que nos permitirá encontrar un estimador minimax de la media de una distribución normal.

Teorema 2. *Supongamos que $\delta(\mathbf{X})$ sea un estimador tal que (i) $R(\delta, \boldsymbol{\theta}) = C\ \forall \boldsymbol{\theta}$, (ii) existe una sucesión de distribuciones a priori γ_k tales que*

$$\lim_{k\to\infty} r(\delta_{\gamma_k}, \gamma_k) = C.$$

Luego δ es minimax.

Demostración: Supongamos que δ no es minimax. Luego existe otro estimador δ' tal que

$$R(\delta', \boldsymbol{\theta}) < \sup_{\boldsymbol{\theta}} R(\delta, \boldsymbol{\theta}) = C \ \ \forall \boldsymbol{\theta}. \qquad (10.32)$$

Integrando ambos miembros de (10.32), se tiene que

$$r(\delta', \gamma_k) = \int ... \int R(\delta', \boldsymbol{\theta})\gamma_k(\boldsymbol{\theta})d\boldsymbol{\theta} < C \int ... \int \gamma_k(\boldsymbol{\theta})d\boldsymbol{\theta} = C.$$

Luego de acuerdo (ii), existe n^* tal que

$$r(\delta', \gamma_{n^*}) < r(\delta_{\gamma_{n^*}}, \gamma_{n^*}),$$

contradiciendo el hecho de que $\delta_{\gamma_k^*}$ es la solución de Bayes correspondiente a γ_{k^*}.

Ejemplo 4. Consideremos una muestra aleatoria $\mathbf{X} = (X_1, ..., X_n)$ de una distribución N$((\theta, \sigma^2)$, donde σ^2 conocida. El estimador $\delta(\mathbf{X}) = \bar{X}$ tiene como funcion de riesgo $R(\delta, \theta) = \sigma^2/n$, y por lo tanto se cumple la condición (i) del Teorema 2. Por otro lado consideremos una sucesión de distribuciones a priori γ_k=N$(0, \tau_k^2)$ con $\tau_k^2 \to \infty$. Usando una función de pérdida cuadrática, de acuerdo a lo visto en el ejemplo 2, los estimadores de Bayes son

$$\delta_{\gamma_k} = w_k\bar{X},$$

donde

$$w_k = \frac{n/\sigma^2}{(n/\sigma^2) + (1/\tau_k^2)}. \tag{10.33}$$

Es fácil ver que

$$\lim_{k\to\infty} w_k = 1 \tag{10.34}$$

y que

$$\lim_{k\to\infty} \tau_k^2 \,(1 - w_k)^2 = \lim_{k\to\infty} \tau_k^2 \frac{1/\tau_k^4}{((n/\sigma^2) + (1/\tau_k^2))} = 0 \tag{10.35}$$

Por otro lado se tiene

$$R(\delta_{\gamma_k}, \theta) = var_\theta(\delta_{\gamma_k}) + (\theta - E_\theta(\delta_{\gamma_k}))^2 = w_k^2\frac{\sigma^2}{n} + (1 - w_k)^2\theta^2.$$

Luego

$$r(\delta_{\gamma_k}, \gamma_k) = E_\gamma(R(\delta_{\gamma_k}, \boldsymbol{\theta})) = w_k^2\frac{\sigma^2}{n} + (1 - w_k)^2\tau_k^2.$$

Luego de acuerdo (10.34) y a (10.35) se tiene que

$$\lim_{k\to\infty} r(\delta_{\gamma_k}, \gamma_k) = \frac{\sigma^2}{n}$$

Por lo tanto se cumple la condición (ii) del Teorema 2, y el estimador $\delta(\mathbf{X}) = \bar{X}/n$ es minimax.