Tarea n°4

Nos interesamos a la teoría general de la M-estimación. En lo que sigue, supongamos dado una muestra $X_1, \ldots, X_n \in \mathcal{X}$ de datos de distribución desconocida P. Notamos P_n la medida empírica asociada. Sea Θ un espacio (un subespacio de un espacio métrico) que llamamos espacio de parámetros (posiblemente de dimensión infinita). Supongamos dada una función de perdida $l : \Theta \times \mathcal{X} \to \mathbb{R}$. Denotaremos $l_\theta(X)$ una valuación de la función l al punto (θ, X) de manera que l_θ sea una función $\mathcal{X} \to \mathbb{R}$. El M-estimador definido sobre la función de perdida l se define como

$$\hat{\theta}_n = \arg\max_{\theta \in \Theta} P_n[l_\theta].$$

Supongamos que θ_0 es un parámetro de Θ único tal que

$$\theta_0 = \arg\max_{\theta \in \Theta} P[l_\theta]$$

P.1 Mostrar que la media empírica y la mediana empírica son M-estimadores.

P.2 Supongamos que $\theta \mapsto P[l_\theta]$ es un mapeo continuo y que $\mathcal{F} = \{l_\theta : \theta \in \Theta\}$ es P-Glivenko-Cantelli. Mostrar que $\hat{\theta}_n$ es un estimador consistente de θ_0.

P.3 Mostrar que si el conjunto de parámetros Θ es compacto (al respecto de una distancia d), que el mapeo $\theta \mapsto P[l_\theta]$ es continuo y que el conjunto $\mathcal{F}$ definido en el problema anterior tiene una envolvente $F \in L_1(\mathcal{X})$, entonces el conjunto $\mathcal{F}$ es P-Glivenko-Cantelli.

P.4 Supongamos que $\Theta \subset \mathbb{R}^k$ es un conjunto convexo y que $\theta \mapsto P[l_\theta]$ es un mapeo continuo y convexo. Se supone que por un $\varepsilon > 0$ dado F_ε dado por

$$F_\varepsilon = \sup_{\|\theta - \theta_0\| \leq \varepsilon} |l_\theta|$$

es tal que $F_\varepsilon \in L_1(\mathcal{X})$. Usando P.2 y P.3, mostrar que el pseudo-estimador $\tilde{\theta}_n = \alpha \hat{\theta}_n + (1-\alpha)\theta_0$ donde $\alpha = \frac{\varepsilon}{\varepsilon + \|\hat{\theta}_n - \theta_0\|}$ es consistente. (**Hint :** Ver que se puede reducir al conjunto $\tilde{\Theta} = \{\theta \in \Theta : \|\theta - \theta_0\| \leq \varepsilon\}$). Concluir que $\hat{\theta}_n$ es consistente.

P.5 (Ejemplo) Supongamos que las variables X_i toman la forma $X_i = (Y_i, Z_i)$ donde $Y_i \in \{0, 1\}$ y $Z_i \in \mathbb{R}$ una covariable. Supongamos que la muestra proviene de una distribución logística dada por

$$P_{\theta_0}(Y = 1 | Z = z) = \frac{1}{1 + \exp(\alpha_0 + \beta_0 z)}.$$

Mostrar, usando lo anterior, que el estimador MLE (máximo de verosimilitud) de $\theta_0 = (\alpha_0, \beta_0) \in \mathbb{R}^2$ es un estimador consistente.

P.6 (Normalidad asintótica) $\Theta \subset \mathbb{R}$ y θ_0 pertenece al interior de Θ. Supongamos que $\hat{\theta}_n$ es consistente. Supongamos que existe $\varepsilon > 0$ dado tal que

1. $\forall \theta$ tal que $|\theta - \theta_0| < \varepsilon$, $\forall x \in \mathcal{X}$, $\theta \mapsto l_\theta(x)$ es derivable de derivada $g_\theta(x) = \partial_\theta l_\theta(x)$.
2. $\mathcal{F}_\varepsilon = \{g_\theta : |\theta - \theta_0| < \varepsilon\}$ es P-Donsker y tal que la envolvente $F_\varepsilon \in L_2(\mathcal{X})$.
3. Por $\theta \to \theta_0$,
$$P(g_\theta - g_{\theta_0}) = \sigma(\theta - \theta_0) + o(|\theta - \theta_0|)$$
por un $\sigma > 0$.
4. Se cumple que $P(g_\theta - g_{\theta_0})^2 \to 0$ cuando $\theta \to \theta_0$ y denotamos $J = P(g_{\theta_0}^2)$.

Mostrar que $\hat{\theta}_n$ cumple un teorema del limite central de varianza que se dará en función de σ y J. (**Hint :** Mostrar que $P_n(g_{\hat{\theta}_n}) = 0$ y usar una descomposición de este termino que involucra el termino $(P_n - P)(g_{\hat{\theta}_n} - g_{\theta_0}) = o_P(n^{-1/2})$)