Introduction

The following are the standard regularity assumptions generally needed in statistics. Uses of these conditions are made in theorem and result statements found in other documents[1]. These conditions were developed to support Fisher's Maximum (ML) Likelihood methodologies (1921, 1925ff), who, according to Norden [1], never gave them in his proofs of MLE efficiency. They were developed later by Hotelling (1931), Doob (1934), Cramer (1946), Rao (1946), Wald (1949), Basu (1952), Le Cam (1953), and others, and were fairly well established by the early 1950's. They were needed in support of the existence, uniqueness and asymptotic efficiency of the MLE. Subsequent conditions were established to handle problems such as inconsistent or superefficient MLE's, etc.

We will generally assume $X_1, X_2, \cdots X_n$ is an i.i.d. set (sample) of random variables defined on the underlying probability space $(\Omega, \mathcal{F}, P)$ in which $X_i \in R^d$ has parametric d.f. F_X with density $f(x \mid \theta)$ w.r.t a σ-finite measure μ. For convergence results we assume an i.i.d. sequence $\{X_i\}_{i=1}^{\infty}$ of such random variables. Some versions of the conditions are stated with *d=1*, $(X_i \in R)$. Broad specifications such as $E(X^k) < \infty$ are invoked by problem statement. Note that higher order moment conditions on random vectors require tensor analysis for *k*>2 and are not generally addressed in this document.

Regularity Conditions for Interchanging limits and integrals/sums

0.1 A Lipschitz condition on bounded variation of $f(x \mid \theta)$, $\ddot{f}(x \mid \theta)$, etc., is needed for LDC to exchange limits and integrals. Let $f(x \mid \theta)$ be differentiable at $\theta = \theta_0$, i.e., for $\delta \in R^k$, $\|\delta\| \cdot T(x,\theta)\big|_{\theta=\theta_0} \doteq \|\delta\| \cdot \nabla f(x,\theta)\big|_{\theta=\theta_0} = f(x,\theta_0+\delta) - f(x,\theta_0) + o(\|\delta\|)$ holds for every x, and let there be a scalar function $g(x \mid \theta_0)$ and constant $\delta_0 \in R_+$ such that

(i) $\| f(x,\theta_0+\delta) - f(x,\theta_0)\| \le \|\delta\| \cdot g(x \mid \theta_0)$ for all x and $\|\delta\| \le \delta_0$, and

(ii) $\int g(x \mid \theta_0) dx < \infty$

0.2 Let $f(x \mid \theta)$ be differentiable in $\theta \in \Theta$, $\exists \delta_0 \in R_+$ and there be a scalar function $g(x \mid \theta) \ni$

(i) $\int g(x \mid \theta) dx < \infty,\ \forall \theta \in \Theta$ and (ii) $\left\| \frac{\partial}{\partial \theta} f(x \mid \theta) \big|_{\theta=\theta'} \right\| \le g(x \mid \theta)\ \forall\ \theta' \in \Theta \ni \|\theta' - \theta\| \le \delta_0$. The latter is the Lipschitz condition extended by the MVT to provide a uniform result in $\theta \in \Theta$. Note that δ_0 is implicitly a function of θ.

0.2′ (Le Cam). Same as 0.2, except $g(x \mid \theta) > 0$, $E(g(x \mid \theta)) < \infty$, and we have that

$\sup_{\theta' : \|\theta'-\theta\| \le \delta_0} \left\| \frac{\partial^2 \log f(x \mid \theta)}{\partial \theta' \partial \theta'^T} \right\|_F \le g(x \mid \theta)$ a.e. Here the Frobenius norm $\|A\|_F = \sqrt{\Sigma \Sigma |a_{ij}|^2}$.

[1] See for example J. Dobelman, Handy List of Statistical Limit Results, Rice University, Houston, TX, 2012.

Classical Statistical Regularity Conditions

1. Identifiability of θ: $f(x\,|\,\theta)$ is distinct $\forall\,\theta \in \Theta$; i.e., $\theta_1 \neq \theta_2$, $f(x\,|\,\theta_1) \neq f(x\,|\,\theta_2)$. This condition is sometimes stated in terms of the Kullback-Leibler distance between densities f_θ, where for

$\theta_1 \neq \theta_2, D(f_1, f_2) = \int f(x\,|\,\theta_1)\log\left(\dfrac{f(x\,|\,\theta_1)}{f(x\,|\,\theta_2)}\right)dx > 0$.

2. Common support: $\{x : f(x\,|\,\theta) > 0\}$ is the same $\forall\,\theta \in \Theta$; i.e., the support is independent of θ.

3. $\Theta \subset R^k$ is an open set (rectangle), not necessarily finite [finite, semi-infinite, or infinite].

   3′. θ_0 is an interior point of W, an open set contained in $\Theta \subset R^k$

4. $f(x\,|\,\theta)$ is differentiable in $\theta \in \Theta \quad \forall x \in R^d \quad \mu -$ a.e.

   4′. $f(x\,|\,\theta)$ is differentiable in $\theta \in W \subset \Theta$, $\forall x \in R^d \quad \mu -$ a.e.

   4″. $f(x\,|\,\theta)$ is twice differentiable in $\theta \in \Theta \;\; \forall x \in \text{Supt}(f) \;\; \mu -$ a.e. , with $\dfrac{\partial^2}{\partial\theta^2}\iint \rightleftharpoons \iint \dfrac{\partial^2}{\partial\theta^2}$

   4‴. $f(x\,|\,\theta)$ is 3 times differentiable in $\theta \in \Theta \;\; \forall x \in \text{Supt}(f) \;\; \mu -$ a.e. , with $\dfrac{\partial^3}{\partial\theta^3}\iiint \rightleftharpoons \iiint \dfrac{\partial^3}{\partial\theta^3}$

5. Unique MLE. The likelihood equation (LE), $\dfrac{\partial}{\partial\theta}L(\theta\,|\,x) = 0$, has a single, unique root (RLE) $\forall n,\; x \in R^d$.

   5′. There exists a unique MLE $\hat{\theta}_n$ for each *n*, and for all $\theta \in \Theta$.

6. (Expected) Fisher Information (FI) defined, positive, finite: $0 < I(\theta) < \infty$, where $I(\theta) = E\left[\left(\dfrac{\partial \ln f(x\,|\,\theta)}{\partial\theta}\right)^2\right]$ and is often denoted $I(\theta_0)$ where θ_0 is the true parameter. In the multiparameter case, FI is positive definite, with $I(\theta) = E\left[\nabla\ell\nabla\ell^T\right]$. When 4″ is in force, $I(\theta)$ is also equal to $-E\left[\nabla^2\ell\right] = -E\left[\nabla\nabla\ell^T\right] = -E[\mathrm{H}_\ell(\theta)]$, and is equal to $E[J(\theta)]$, where *J* is the observed FI, although *J* is normally evaluated at $\hat{\theta}_{ML}$ and not averaged over x.

7. Boundedness of variation of second derivative of LF: For each $x \in \text{Supt}(f)$, $\exists d \in R_+$ and function $B(x) \ni \left|\dfrac{\partial^2}{\partial\theta^2}\ln f(x\,|\,\theta)\right| \leq B(x)$ with $\theta \in [\theta_0 - d,\, \theta_0 + d]$ and $E[B(x)] < \infty$. Both B and d can be functions of θ.

7′. In addition to 7, $B \neq B(\theta)$, and there are integrable $F_1(x)$ and $F_2(x)$ such that $\left|\frac{\partial f}{\partial \theta}\right| < F_1(x)$ and $\left|\frac{\partial^2 f}{\partial \theta^2}\right| < F_2(x)$

Extended Regularity Conditions

References

1. Norden
2. Lehmann & Casella
3. Casella & Berger
4. D.Cox
5. Stuart & Ord
6. Polansky
7. Young & Smith
8. Gentle
9. Etc.