STA3000: Likelihood Formalities

The likelihood function is defined as (proportional to) the density function, and this is a density with respect to some dominating measure. Since θ varies in Θ, we need f to be a density function with respect to the same dominating measure for each value of θ. Schervish (p.13) states it this way:

Let $(S, \mathcal{A}, \mu)$ be a probability space, and let $(\mathcal{X}, \mathcal{B})$ be a Borel space. Let $X : S \longrightarrow \mathcal{X}$ be measurable. The parametric family of distributions for X is the set

$$\{P_\theta : \forall A \in \mathcal{B}, P_\theta(A) = \Pr(X \in A), \theta \in \Theta\}.$$

Assume that each P_θ, considered as a measure on $(\mathcal{X}, \mathcal{B})$ is absolutely continuous with respect to a measure ν on $(\mathcal{X}, \mathcal{B})$. We write

$$f(x;\theta) = \frac{dP_\theta}{d\nu}(x);$$

this is the likelihood function for θ.

Some books describe the likelihood function as the Radon-Nikodym derivative of the probability measure with respect to a dominating measure. Sometimes the dominating measure is taken to be P_{θ_0} for a fixed value $\theta_0 \in \Theta$. When we consider probability spaces and/or parameter spaces that are infinite dimensional, it is not obvious what to use as a dominating measure. For counting processes, this is done rigorously in Ch.II of Andersen et al. The result is Jacod's formula for the likelihood ratio:

Suppose we have a counting process $N(\cdot)$ on $[0, \tau]$, and a filtration $\mathcal{F}_t = \mathcal{F}_0 \cup \sigma\{N(s); s \le t\}$, with $\mathcal{F} = \mathcal{F}_\tau$. A counting process is a piecewise constant, non-decreasing, stochastic process with jumps of size $+1$. It can be shown to be a local submartingale, with compensator Λ. Suppose P and $\tilde{P}$ are two probability measures on $\mathcal{F}$, for which the two compensators are Λ and $\tilde{\Lambda}$. Suppose $\tilde{P}$ is absolutely continuous with respect to P. If Λ and $\tilde{\Lambda}$ are absolutely continuous a.s. P, then

$$\frac{d\tilde{P}}{dP} = \left.\frac{d\tilde{P}}{dP}\right|_{\mathcal{F}_0} \frac{\prod_t \tilde{\lambda}(t)^{\triangle N(t)} \exp\{-\tilde{\Lambda}(\tau)\}}{\prod_t \lambda(t)^{\triangle N(t)} \exp\{-\Lambda(\tau)\}}.$$

Except for the somewhat unfamiliar notation, this is identical to the likelihood function for the non-homogeneous Poisson process we discussed in L2 (following SM, §6.5), which was expressed

$$\prod_{j=1}^n \lambda(t_j) \exp\{-\int_0^\tau \lambda(u)du\}, \quad 0 < t_1 < \cdots < t_n < \tau.$$

References

Andersen, P.K., Borgan, O., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer, New York.
[BNC] Barndorff-Nielsen, O.E. and Cox, D.R. (1994). *Inference and Asymptotics.* Chapman & Hall, London.
Schervish, M.J. (1995). *Theory of Statistics.* Springer, New York.
[SM] Davison, A.C. (2003). *Statistical Models.* Cambridge University Press, Cambridge.