3.1 The Likelihood Principle

a random variable with density f_θ. The outcome of E is $X = x$. The inferential evidence about θ from (E, x) is written $Ev(E, x)$.

Weak Sufficiency Principle: Consider an experiment E and suppose that $T(x)$ is sufficient for θ. Then if $T(x_1) = T(x_2)$ we have $Ev(E, x_1) = Ev(E, x_2)$.

Definition 3.1.1: A statistic S is ancillary for θ if its distribution is free of θ.

Definition 3.1.2: A family $(f_\theta, \theta \in THETA)$ of distributions is boundedly complete if for every bounded real function g such that $\int g(x) f_\theta(x) dx = 0$ for every θ, we have that $g \equiv 0$ a.s.

Theorem 3.1.1: (Basu's theorem)
Suppose T is sufficient and U is ancillary. If (f_θ) is boundedly complete, then T and U are independent.

Weak Conditionality Principle: Let $E_1 = (X_1, \theta, f_\theta^1)$ and $E_2 = (X_2, \theta, f_\theta^2)$ where the unknown parameter θ is common to the two experiments. Consider the mixed experiment $E*$ wherein J=1 or 2 is observed, each with probability $\frac{1}{2}$, and experiment E_J is performed. Then $E* = (X*, \theta, f_\theta^*)$ where $X* = (J, X_J)$ and $f_\theta^*(j, x_j) = \frac{1}{2} f_\theta^j(x_j)$. Then $Ev(E*, (j, x_j)) = ev(E_j, x_{j)}$.

Likelihood Principle: Let E_1, E_2 be experiments with the same quantity θ in each. Suppose that for particular outcomes x_1, x_2 from the two experiments we have $L_{x_1}(\theta) = cL_{x_2}(\theta)$. Then $Ev(E_1, x_1) = Ev(E_2, x_2)$. In particular, $Ev(E, x)$ depends on E and x only through the likelihood $L_x(\theta)$.

Theorem 3.1.2: (Birnbaum's theorem)
(WCP and WSP) is equivalent to LP.

3.2 Minimal sufficiency

Definition 3.2.1: A sufficient statistic T is minimal sufficient if for any sufficient statistic U there is a function h so that $T = h(U)$.

Theorem 3.2.1:
(a) The likelihood ratio $f(x;\theta)/f(x;\theta_0)$ is a function of $T(x)$ iff T is sufficient.

(b) Let $\mathbf{P}$ be a finite family with densities $f_i, i = 0, \ldots, k$. Then

$$U(X) = \left(\frac{fsub1(X)}{f_0(X)}, \ldots, \frac{f_k(X)}{f_0(X)} \right)$$

is minimal sufficient.

(c) Let $\mathbf{P}$ be a family of distributions with common support, and let $\mathbf{P}_0 \subset \mathbf{P}$. If T is minimal sufficient in $\mathbf{P}_0$ and sufficient in $\mathbf{P}$, then T is minimal sufficient in $\mathbf{P}$.

3.3 Barndorff-Nielsen's formula

Theorem 3.3.1: (Barndorff-Nielsen's formula):
Suppose the minimal sufficient statistic is $(\hat{\theta}, a)$ where a is ancillary for θ. Then

$$f(\hat{\theta}|a;\theta) \doteq c\left\{\frac{L(\theta;\hat{\theta},a)}{L(\hat{\theta};\hat{\theta},a}\right\}|j(\hat{\theta})|^{\frac{1}{2}} = p*(\hat{\theta}|a)$$

where $L(\theta; x) = f(x; \theta)$.

Definition 3.3.1: Consider $\theta = (\psi, \lambda)$ where λ is a nuisance parameter. The profile likelihood of ψ is $L(\psi, \hat{\lambda}^{\psi})$ where $\hat{\lambda}_{\psi}$ solves the likelihood equation for lambda at a given value of ψ.

Consider the class of distributions with minimal sufficient statistic s having the factorization

$$f(s;\psi,\lambda) = f(s_1|s_2;\psi)f(s_2;\psi,\lambda).$$

A saddlepoint approximation to the conditional log likelihood $l_c(psi) = \log f(s_1|s_2;\psi)$ is then given by the **adjusted profile likelihood**

$$l_c(\psi) \doteq l(\psi,\hat{\lambda}_{\psi}) + \tfrac{1}{2}\log|J_{\lambda\lambda}(\psi,\hat{\lambda}_{\psi})|.$$

The Barndorff-Nielsen formula applies to approximate the density of the conditional mle $\hat{\psi}_c$, namely

$$f(\hat{\psi}_c;\psi) \doteq c[-l_c^{''}(\hat{\psi}_c)]^{\frac{1}{2}}\,\frac{L_c(\hat{\psi}_c)}{L_c(\psi)}\,.$$