CHAPTER 7. TAIL PROBABILITIES

In exponential families the probability under θ of a set generally falls off exponentially fast as the distance of the set from $\xi(\theta)$ increases. This section contains several results of this form. The first of these will be improved later, but it is included here because of its simplicity of statement and proof.

Throughout this chapter let $\{p_\theta\}$ be a steep canonical exponential family. (Most of the results hold with possibly minor modifications for non-minimal families, and many also hold for non-steep families.)

FIXED PARAMETER (Via Chebyshev's Inequality)

7.1 Theorem

Fix $\theta_0 \in N^\circ$. Choose ε so that $\{\theta\colon ||\theta - \theta_0|| \le \varepsilon\} \subset N^\circ$. Then there exists a constant $c < \infty$, such that

$$\Pr_{\theta_0} H^+(v, \alpha) \quad < \quad c \exp(-\varepsilon\alpha) \tag{1}$$

for all $v \in R^k$ with $||v|| = 1$ and all $\alpha \in R$.

Proof. Let

$$c \;=\; \exp(\sup\{\psi(\theta) - \psi(\theta_0)\colon\; ||\theta - \theta_0|| = \varepsilon\}) \tag{2}$$

and let $\theta_\varepsilon = \theta_0 + \varepsilon v$. Then

$$\begin{aligned}
\Pr_{\theta_0}\{H^+(v,\alpha)\} &= \int_{H^+(v,\alpha)} \exp(\theta_0 \cdot x - \psi(\theta_0))\nu(dx) \\
&= \int_{H^+(v,\alpha)} \exp(\theta_0 \cdot x + (\varepsilon v) \cdot x - (\varepsilon v) \cdot x - \psi(\theta_0))\nu(dx) \\
&\leq \left(\int_{H^+(v,\alpha)} \exp(\theta_\varepsilon \cdot x - \psi(\theta_\varepsilon))\nu(dx)\right)\exp(\psi(\theta_\varepsilon) - \psi(\theta_0) - \varepsilon\alpha) \\
&\leq c\exp(-\varepsilon\alpha) \quad . \qquad ||
\end{aligned}$$

Note that (2) provides a specific formula for the constant appearing in (1).

In specific situations the bound provided in Theorem 7.1 can be improved in various ways. However the following converse result shows that Theorem 7.1 always comes within an arbitrarily small amount of yielding the best exponential rate of decrease for tail probabilities.

7.2 Proposition

Let $\theta_0 \in N^\circ$. Suppose there exists a $c < \infty$ and $\varepsilon > 0$ such that 7.1(1) is valid for all $v \in R^k$ with $||v|| = 1$ and all $\alpha > 0$. Then $\{\theta: ||\theta - \theta_0|| < \varepsilon\} \subset N^\circ$.

(Thus, if for some $\varepsilon > 0$, $c < \infty$, a bound of the form 7.1(1) is valid for all v with $||v|| = 1$ and all $\alpha > 0$, then Theorem 7.1 will verify such a bound for any $\varepsilon' < \varepsilon$.)

Proof. We leave the proof as an exercise. ||

When $\varepsilon = \inf \{||\theta - \theta_0|| : \theta \notin N\}$ then 7.1(1) may or may not be valid for all α, ν. The following example demonstrates this.

7.3 Example

Relative to Lebesgue measure, let

(1) $$f_{\sigma,k}(y) = \Gamma(k)y^{k-1}\, e^{-y/\eta}/\eta^k \qquad y > 0$$

$$0 \qquad y \le 0 \quad .$$

This is the gamma density with scale parameter η and shape parameter k. Let $x_1 = y$, $x_2 = \ln y$, $\theta_1 = -1/\eta$, $\theta_2 = (k - 1)$, and let ν be the measure induced by the map $y \to x$ when y has Lebesgue measure on $(0, \infty)$. One then has a standard exponential family of order 2 with

$$\psi(\theta) = (\theta_2 + 1)\ln(-\theta_1) - \ln\Gamma(\theta_2 + 1)$$

and

(2) $$N = (-\infty, 0) \times (-1, \infty), \qquad K = \{(x_1, x_2):\ x_1 \ge 0,\ \ x_2 \ge \ln x_1\} \quad .$$

When $k = 1$ (i.e. $\theta_2 = 0$) the resulting one-parameter exponential family is that of exponential distributions with intensity $|\theta_1|$. For this family

$$\Pr_{\theta_1=-1}\{x_1 > \alpha\} = e^{-\alpha} \qquad \text{for all} \qquad \alpha > 0$$

so that 7.1 holds with $v = 1$ and $\varepsilon = 1 = \inf\{||\theta - \theta_0|| : \theta \notin N\}$. On the other hand, for $\theta_2 = 1$ the resulting one-parameter gamma family has

$$\Pr_{\theta_1=-1}\{x_1 > \alpha\} = (\alpha + 1)e^{-\alpha} \qquad \text{for all} \qquad \alpha > 0.$$

Thus here 7.1(1) fails to hold when $v = 1$ and $\varepsilon = 1 = \inf\{||\theta - \theta_0|| : \theta \notin N\}$

When $N = R^k$ Theorem 7.1 says only that $\Pr_{\theta_0}\{H^+(u, \alpha)\} = O(e^{-k\alpha})$ for all $k > 0$. However, much smaller bounds may be valid for these tail probabilities. Consider for example the following well known facts:

(3) $$\int_\alpha^\infty e^{-t^2/2}\, dt \le e^{-\alpha^2/2}/\alpha \qquad \text{for} \qquad \alpha > 0$$

and

(4) $$\int_{\alpha}^{\infty} e^{-t^2/2}\,dt \sim e^{-\alpha^2/2}/\alpha \quad \text{as} \quad \alpha \to \infty \quad .$$

Thus, suppose X is normal, mean 0, variance 1. Then, from (3)

(5) $$\Pr\{X > \alpha\} \leq e^{-\alpha^2/2}/\alpha(2\pi)^{\frac{1}{2}} \quad \text{for} \quad \alpha > 0 \quad .$$

It can be seen from (4) that this bound is asymptotically accurate as $\alpha \to \infty$.

Theorem 7.5 contains a bound which easily yields the statement

(6) $$\Pr\{X > \alpha\} \leq e^{-\alpha^2/2}$$

for this situation. This is much better than what is available from 7.1(1) but is still inferior to (5).

Theorem 7.1 applies to probabilities of large deviations defined by half spaces but can easily be converted to a statement about any shape of set, as follows.

7.4 Corollary

Consider a standard exponential family. Fix $\theta_0 \in N^\circ$. Let $x_0 \in R^k$. Let S be any set. Let $\rho = \inf\{||x - x_0|| : x \notin S\}$, and define ε as in Theorem 7.1. Then there is a $c < \infty$ such that

(1) $$P_{\theta_0}(\{(X - x_0)/\alpha \notin S\}) < c\exp(-\varepsilon\rho\alpha) \quad \text{for all} \quad \alpha \in R \quad .$$

Proof. It suffices to prove the corollary for $x_0 = 0$ and S the open sphere of radius ρ about the origin.

There exists $\rho' < \rho$ and $\varepsilon' < \inf\{||\theta - \theta_0|| : \theta \notin N\}$ such that $\varepsilon'\rho' = \varepsilon\rho$. There exists a finite set of unit vectors $\{a_i: i=1,\ldots,n\}$ such that $\bigcap_{i=1}^{n} \{x: x \cdot a_i < \rho'\} \subset S$. Thus $\Pr_{\theta_0}\{X/\alpha \notin S\} < \sum_{i=1}^{n} \Pr_{\theta_0}\{X \cdot a_i > \alpha\rho'\}$

$\leq \sum_{i=1}^{n} c_i \exp(-\alpha\rho'\varepsilon') \leq c\exp(-\varepsilon\rho\alpha)$ by Theorem 7.1 where $c < \infty$ is an appropriate constant. ||

FIXED PARAMETER (Via Kullback-Leibler Information)

It is possible to use the Kullback-Leibler information number (i.e. entropy) to improve the preceding bound. See the exercises for some applications of this bound to asymptotic theory.

7.5 Theorem

Let $\theta_0 \in N°$ and $\bar{H}^+ = \bar{H}^+(v, \alpha)$. Then

(1) $$P_{\theta_0}(\bar{H}^+) \leq \exp(-\tilde{K}(\bar{H}^+, \xi(\theta_0))) \quad .$$

Proof. Suppose first that

(2) $$\bar{H}^+ \cap K° \neq \phi \quad .$$

Let $\tilde{\xi} = \tilde{\xi}_{\bar{H}^+}(\theta_0)$. Note that $\xi \in \bar{H}^+ \cap K°$ by Theorem 6.13. Hence $\tilde{\theta} = \theta(\tilde{\xi}) \in N°$. (This is precisely the situation pictured in Figure 6.14(1).) Now,

(3) $$\tilde{k} = \tilde{K}(\bar{H}^+, \xi(\theta_0)) = (\tilde{\theta} - \theta_0) \cdot \tilde{\xi} - \psi(\tilde{\theta}) + \psi(\theta_0)$$

$$\leq (\tilde{\theta} - \theta_0) \cdot x - \psi(\tilde{\theta}) + \psi(\theta_0) \qquad \forall \ x \in \bar{H}^+$$

by definition and by 6.13(2). This yields

$$P_{\theta_0}(\bar{H}^+) = \int_{\bar{H}^+} p_{\theta_0}(x)\nu(dx) = \int_{\bar{H}^+} \frac{p_{\theta_0}(x)}{p_{\tilde{\theta}}(x)} p_{\tilde{\theta}}(x)\nu(dx)$$

$$= \int_{\bar{H}^+} \exp((\theta_0 - \tilde{\theta}) \cdot x - \psi(\theta_0) + \psi(\tilde{\theta}))p_{\tilde{\theta}}(x)\nu(dx)$$

$$\leq \int_{\bar{H}^+} \exp(-\tilde{k})p_{\tilde{\theta}}(x)\nu(dx) \leq e^{-\tilde{k}} \quad ,$$

which is the desired result.

Now suppose $\bar{H}^+ \cap K \neq \phi$ but $\bar{H}^+ \cap K° = \phi$. Then

(4) $$\lim_{\varepsilon\downarrow 0} \tilde{K}(\bar{H}^+(v, \alpha - \varepsilon), \xi(\theta_0)) = \tilde{K}(\bar{H}^+(v,\alpha), \xi(\theta_0)) \le \infty$$

since $\tilde{K}(\cdot, \xi(\theta_0))$ is lower semi-continuous (by definition), satisfies

$$\lim_{||\xi||\to\infty} \tilde{K}(\xi, \xi(\theta_0)) = \infty$$

(by 6.5(5)), and since $\tilde{K}(\bar{H}^+(v, \alpha), \xi(\theta_0)) \ge \tilde{K}(\bar{H}^+(v, \alpha - \varepsilon), \xi(\theta_0))$ for all $\varepsilon > 0$. Hence

(5) $$P_{\theta_0}(\bar{H}^+) = \lim_{\varepsilon\downarrow 0} P_{\theta_0}(\bar{H}^+(v, \alpha-\varepsilon)) \le \lim_{\varepsilon\downarrow 0} \exp(-\tilde{K}(\bar{H}^+(v, \alpha-\varepsilon), \xi(\theta_0)))$$

$$= \exp(-\tilde{K}(\bar{H}^+, \xi(\theta_0))) \quad . \qquad ||$$

(We leave as an exercise to verify that

(6) $$\tilde{K}(\bar{H}^+, \xi(\theta_0)) = \infty \quad \text{if and only if} \quad P_{\theta_0}(\bar{H}^+) = 0 \quad .)$$

Note that the Kullback-Leibler information enters into the above only as a convenient way of identifying the sup $\{(\tilde{\theta} - \theta_0) \cdot x - \psi(\tilde{\theta}) + \psi(\theta_0)$: $x \in H^+\}$. Various other interpretations of K, such as the probabilistic Definition 6.1, do not enter into the above argument.

The connection between Theorem 7.5 and 7.1 is provided by the following lemma.

<u>7.6 Lemma</u>

Let $\theta_0 \in N°$ and $H^+ = H^+(v, \alpha)$. Suppose $\theta = \theta_0 + \varepsilon v \in N°$. Then

(1) $$\tilde{K}(H^+, \xi(\theta_0)) \ge \psi(\theta_0) - \psi(\theta) + \varepsilon\alpha \quad .$$

Proof. Let $\tilde{\xi} = \tilde{\xi}_{\bar{H}^+}(\theta_0)$ as in Theorem 7.5. Then

$$\tilde{K}(H^+, \xi(\theta_0)) = (\tilde{\theta} - \theta_0) \cdot \tilde{\xi} + \psi(\theta_0) - \psi(\tilde{\theta})$$

$$\ge (\theta - \theta_0) \cdot \tilde{\xi} + \psi(\theta_0) - \psi(\theta)$$

since $\tilde{\theta} = \theta(\tilde{\xi}) = \hat{\theta}_N(\tilde{\xi})$ maximizes $\ell(\cdot, \tilde{\xi})$. Hence

$$\tilde{K}(H^+, \theta_0) \geq \varepsilon v \cdot \tilde{\xi} + \psi(\theta_0) - \psi(\theta) = \varepsilon\alpha + \psi(\theta_0) - \psi(\theta) \ . \quad ||$$

Applying the bound (1) in the formula 7.5(1) yields the earlier formulae, 7.1(1) and (2), of Theorem 7.1.

Note also that in the normal example of Example 7.3, $\tilde{K}(\xi, 0) = \xi^2/2$, and thus 7.5(1) yields 7.3(6).

FIXED REFERENCE SET

The preceding results concern the nature of probabilities of large deviations when the parameter is fixed and the reference set for calculating the probability proceeds to infinity. There is another class of results. These concern the situation when the reference set is fixed and the parameter proceeds to infinity in an appropriate direction. These theorems were exploited in a statistical setting by Birnbaum (1955) and then Stein (1956). Giri (1977) surveys several further applications of this theory.

7.7 Theorem

Let $v \in R^k$, $\alpha \in R$. Let $S_1, S_2 \subset R^k$ with

(1) $$S_2 \subset \bar{H}^-(v, \alpha) \quad ,$$

(2) $$\nu(S_1 \cap H^+(v, \alpha)) > 0 \quad .$$

Let $K \subset N$ be compact. Then there exist constants c and $\varepsilon > 0$ such that

(3) $$\frac{\int_{S_2} e^{\theta \cdot x} \nu(dx)}{\int_{S_1} e^{\theta \cdot x} \nu(dx)} \leq c \exp(-\rho\varepsilon)$$

for all $\theta \in N$ of the form $\theta = \eta + \rho v$ with $\eta \in K$, $\rho > 0$.

Proof. Let $S_1(\varepsilon) = S_1 \cap H^+(v, \alpha + \varepsilon)$. There is an $\varepsilon > 0$ such that $\nu(S_1(\varepsilon)) > \varepsilon > 0$. Then,

$$\frac{\int_{S_2} e^{\theta \cdot x}\, \nu(dx)}{\int_{S_1} e^{\theta \cdot x}\, \nu(dx)} \leq \frac{\int_{S_2} \exp(\rho(v \cdot x - \alpha) + \rho\alpha + \eta \cdot x)\nu(dx)}{\int_{S_1(\varepsilon)} \exp(\rho(v \cdot x - \alpha) + \rho\alpha + \eta \cdot x)\nu(dx)}$$

$$\leq \frac{\int_{S_2} e^{\eta \cdot x}\, \nu(dx)}{e^{\rho\varepsilon} \int_{S_1(\varepsilon)} e^{\eta \cdot x}\, \nu(dx)} \leq c \exp(-\rho\varepsilon)$$

where

$$c = \sup_{\eta \in K} \left(\int_{S_2} e^{\eta \cdot x}\, \nu(dx) \Big/ \int_{S_1(\varepsilon)} e^{\eta \cdot x}\, \nu(dx)\right) < \infty \quad . \tag{4}$$

Here is why $c < \infty$: K is compact and $\nu(S_1(\varepsilon)) > 0$ so that $\inf_{\eta \in K} \int_{S_1(\varepsilon)} e^{\eta \cdot x}\, \nu(dx) > 0$. Also, $\int_{S_2} e^{\eta \cdot x}\, \nu(dx)$ is upper semicontinuous on K by Fatou's lemma, and is finite on K since $K \subset N$. Thus $\sup_{\eta \in K} \int_{S_2} e^{\eta \cdot x}\nu(dx) < \infty$. ||

The preceding theorem really concerns the relationship of probabilities for the sets S_2 and $S_1(0) = S_1 \cap H^+(v, \alpha)$ contained in separate half spaces. Note again the dual relationship, connecting $\theta \in N$ and $H \subset K$ in Theorem 7.7. Because of this relationship it is often revealing in such contexts to superimpose both the sample space and parameter space on a single plot. This is done in Example 7.12(1).

Here are some corollaries to the Theorem, the second of which will be used in the example. The first of these corollaries may be instructively compared to Theorem 7.1.

7.8 Corollary

Let $v \in R^k$, $K \subset N$ be compact, and $S \subset \bar{H}^-(a, \alpha)$. Suppose

(1) $$\nu(H^+(v, \alpha)) > 0 \quad .$$

Then there exist constants c and $\varepsilon > 0$ such that

$$\Pr_\theta(S) \leq c \exp(-\rho\varepsilon)$$

for all $\theta \in N$ of the form $\theta = \eta + \rho v$ with $\eta \in K$, $\rho > 0$. In particular, for any sequence $\{\theta_i \in N: \theta_i = \rho_i v + \eta_i,\ \rho_i \to \infty,\ \eta_i \in K\}$ one has

$$\lim_{i\to\infty} \Pr_{\theta_i}(S) = 0 \quad .$$

Proof. Let $S_2 = H^+(v, \alpha)$. Then by Theorem 7.7

$$\Pr_\theta(S) \leq c \exp(-\rho\varepsilon) \int_{S_2} e^{\theta\cdot x-\psi(\theta)}\,\nu(dx)$$

$$= c \exp(-\rho\varepsilon)\Pr_\theta(S_2) \leq c \exp(-\rho\varepsilon) \quad . \qquad ||$$

7.9 Corollary

Again, let $v \in R^k$, $K \subset N^\circ$ be compact, and $\nu(S) > 0$; and let $\{\theta_i\}$ be any sequence of the form $\theta_i = \rho_i v + \eta_i$ with $\rho_i \to \infty$ and $\eta_i \in K$. Then

(1) $$\lim_{i\to\infty} E_{\theta_i}(v \cdot X) = \sup\{\alpha: \nu(H^+(v, \alpha)) > 0\} \leq \infty \quad .$$

(Note that here we assume $K \subset N^\circ$; not merely $K \subset N$.)

Proof. Let α_0 denote the supremum on the right of (1). Since $E_{\theta_i}(v\cdot X) \leq \alpha_0$ it is only necessary to prove $\liminf_{i\to\infty} E_{\theta_i}(v\cdot X) \geq \alpha_0$. To this end, let $\alpha < \alpha' < \alpha_0$ and $S_2 = H^-(v, \alpha')$. Let $\xi_2(\theta) = E_\theta(X|X \in S_2)$. If $\nu(S_2) = 0$ the result is trivial. Hence, suppose $\nu(S_2) > 0$. Note that $\xi_2(\theta)$ exists and is continuous for all $\theta \in N^\circ$. Hence $\beta = \inf\{v \cdot \xi_2(\eta): \eta \in K\} > -\infty$. Note that

$\beta \leq \alpha'$.

Apply Corollary 2.5 to the conditional exponential family given $X \in S_2$ (generated by $\nu|_{S_2}$) to find

$$E_\theta(v \cdot X | X \in S_2) \geq E_\eta(v \cdot X | X \in S_2) \geq \beta$$

for all $\theta = \eta + \rho v$ with $\rho \geq 0$. Then for such θ,

$$\begin{aligned} E_\theta(v \cdot X) &= Pr_\theta(X \in S_2) \cdot E_\theta(v \cdot X | X \in S_2) \\ &\quad + Pr_\theta(X \notin S_2) \cdot E_\theta(v \cdot X | X \in S - S_2) \\ &\geq (ce^{-\varepsilon\rho}/(1 + ce^{-\varepsilon\rho}))\beta + (1/(1 + ce^{-\varepsilon\rho}))\alpha' \end{aligned}$$

by Theorem 7.7. Hence $E(v \cdot X) > \alpha$ (since $\alpha < \alpha'$) for θ as above for all ρ sufficiently large. This implies $\liminf_{i\to\infty} E_{\theta_i}(v \cdot X) \geq \alpha_0$, since $\alpha < \alpha_0$ was arbitrary. ||

Note the placement of the hyperplane H in the statement of Theorem 7.7. If $S_2 \subset H^-$ and $\nu(S_1 \cap \bar{H}^+) > 0$, but $\nu(S_1 \cap H^+) = 0$, then only a much weaker conclusion is valid. This conclusion is contained in the following corollary.

7.10 Corollary

Let $v \in R^k$, $\alpha \in R$. Suppose

(1) $$S_2 \subset H^-(v, \alpha)$$

and

(2) $$\nu(S_1 \cap \bar{H}^+(v, \alpha)) > 0 \quad .$$

Let $K \subset N$ be compact. Let $\{\theta_i\} \subset N$ be a sequence of the form $\theta_i = \rho_i v + \eta_i$ with $\eta_i \in K$, $\rho_i \to \infty$. Then

(3) $$\lim_{i\to\infty} \frac{P_{\theta_i}(S_2)}{P_{\theta_i}(S_1)} = 0 \quad .$$

Proof. Apply Theorem 7.7 to find

$$\text{(4)} \qquad \lim_{i\to\infty} \frac{P_{\theta_i}(S_2 \cap H^-(v, \alpha - \varepsilon))}{P_{\theta_i}(S_1)} = 0$$

for all $\varepsilon > 0$. Furthermore, if $\rho_i > 0$

$$\text{(5)} \qquad \frac{P_{\theta_i}(S_2 \cap \bar{H}^+(v, \alpha - \varepsilon))}{P_{\theta_i}(S_1)} < \frac{P_{\eta_i}(S_2 \cap \bar{H}^+(v, \alpha - \varepsilon))}{P_{\eta_i}(S_1)} \to 0$$

as $\varepsilon \to 0$ uniformly for $\eta_i \in K$.

(The inequality in (5) follows after applying Corollary 2.23 to the functions $h_c(x) = \chi_{S_1} - c\chi_{S_2 \cap \bar{H}^+(v,\alpha-\varepsilon)}$ with c chosen so that $E_{\eta_i}(h_c(X)) = 0$ to find that $E_{\theta_i}(h_c(X)) \geq E_{\eta_i}(h_c(X))$ for all c and $\rho_i > 0$.) Combining (4) and (5) yields the conclusion of the corollary. ||

7.11 Example

Consider the usual sufficient statistics $\bar{X}$, S^2 derived from a normal (μ, σ^2) sample. As explained in Example 1.2 the statistics $X_1 = \bar{X}$, $X_2 = S^2 + \bar{X}^2$ are the canonical statistics for a two-parameter exponential family with canonical parameters $\theta_1 = n\mu/\sigma^2$, $\theta_2 = -n/2\sigma^2$. Note that $K = \{(x_1, x_2): x_2 \geq x_1^2\}$. For some $c > 1$ consider the conditioning set $Q = \{(x_1, x_2): x_2 \geq cx_1^2\} = \{(\bar{x}, s^2): \bar{x}^2/s^2 \leq 1/(c - 1)\}$. (This is the set on which the usual two-sided t-test (based on $t = \sqrt{n-1}\,\bar{x}/s$) with $n - 1$ degrees of freedom accepts at the appropriate level determined by c.) Fix $\mu = \mu_0$ and let $\sigma^2 \to 0$. Then $(\theta_1, \theta_2) = (n/\sigma^2)(\mu_0, -\tfrac{1}{2})$. Thus (θ_1, θ_2) proceeds down the ray with slope $-\tfrac{1}{2}\mu_0$ as $\sigma^2 \to 0$. Both X and Θ are displayed on the plot in Figure 7.11(1), which shows also K, Q, and this line.

Corollary 7.9 applied to the conditional exponential family given $X \in Q$ (generated by the measure ν restricted to Q) yields

(1) $$\lim_{\sigma^2\to 0} E_{(\mu_0,\sigma^2)}(\mu_0 X_1 - X_2/2|X \in Q)$$

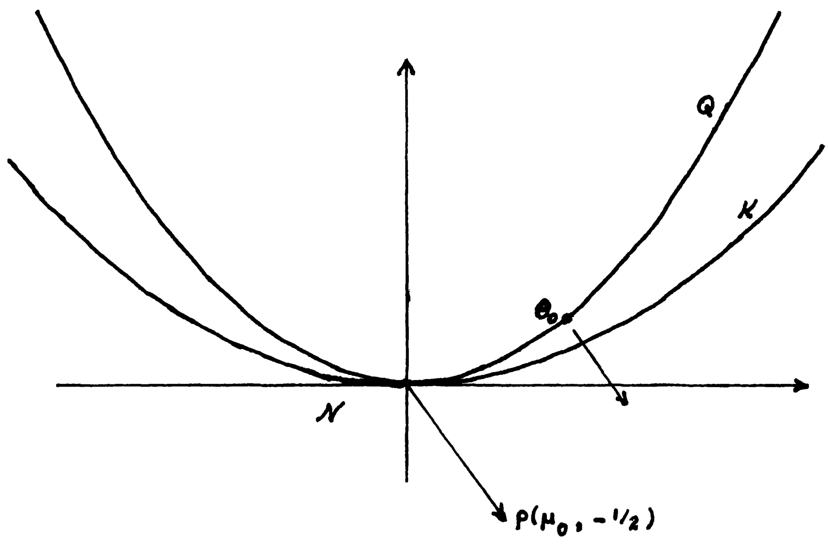
$$= \sup\{\mu_0 x_1 - x_2/2|(x_1, x_2) \in Q\} = \mu_0/2c^2 .$$

Note that $E(\mu_0 X_1 - X_2/2|X \in Q) = (\mu_0, -\frac{1}{2}) \cdot E((X_1, X_2)|X \in Q)$ and that $E((X_1, X_2)|X \in Q) \in Q$. Furthermore since Q is strictly convex $(\mu_0, -\frac{1}{2}) \cdot (x_{1i}, x_{2i}) \to \mu_0/2c^2 = \sup\{(\mu_0, -\frac{1}{2}) \cdot (x_1, x_2): x_1, x_2 \in Q\}$ for a sequence $\{(x_{1i}, x_{2i})\} \subset Q$ if and only if $(x_{1i}, x_{2i}) \to (\mu_0/c, \mu_0^2/c)$. (Note that the tangent to Q at the point $(\mu_0/c, \mu_0^2/c)$ is perpendicular to the ray $(n/\sigma^2)(\mu_0, -\frac{1}{2})$.) Thus

(2) $$\lim_{\sigma^2\to 0} E_{(\mu_0,\sigma^2)}((X_1, X_2)|X \in Q) = (\mu_0/c, \mu_0^2/c) = e_0 \text{ (say)} .$$

In terms of the traditional variables $\bar{X}$, S^2, and $t = \sqrt{n-1}\,\bar{x}/s$ this yields

(3) $$\lim_{\sigma^2\to 0} E_{(\mu_0,\sigma^2)}((\bar{X}, S^2)|\ |t| \le \tau) = \left(\frac{\tau\mu_0}{\tau + n - 1}, \frac{(n-1)\mu_0^2}{(\tau + n - 1)^2}\right)$$



Example 7.11(1): Picture for Example 7.12

COMPLETE CLASS THEOREMS FOR TESTS (Separated Hypotheses)

The preceding results can be used to prove admissibility of many conventional test procedures in univariate and multivariate analysis of variance and in many other testing situations involving exponential families. When combined with the continuity theory for Laplace transforms of Section 2.17 these results yield useful complete class characterizations for certain classes of problems. In many of these cases the characterization precisely describes the minimal complete class. The general theory, as well as a very few specific applications, is described in the remainder of this chapter. Many more applications can be found in the cited references. The results to follow should be compared to the results in the same spirit for estimation which appear in Chapter 4.

7.12 Setting and Definitions

Throughout the remainder of this chapter $\{p_\theta\colon \theta \in \Theta\}$ is a standard exponential family. The parameter space Θ is divided into non-empty null and alternative spaces Θ_0, Θ_1; so that $\Theta = \Theta_0 \cup \Theta_1$. In the customary fashion, a test of Θ_0 versus Θ_1 is uniquely specified by its critical function, ϕ, where $\phi(x) = P(\text{test rejects } \Theta_0 | X = x)$. The power of ϕ is $\pi_\phi(\theta) = E_\theta(\phi)$. A test ϕ_1 is as good as a test ϕ_2 if

$$(1) \qquad \begin{aligned} \pi_{\phi_1}(\theta) &\le \pi_{\phi_2}(\theta) \qquad \theta \in \Theta \\ \pi_{\phi_1}(\theta) &\ge \pi_{\phi_2}(\theta) \qquad \theta \in \Theta \end{aligned}$$

It is better if there is strict inequality for some $\theta \in \Theta$. (Here, and in what follows, we write, "a test ϕ" in place of the more precise but cumbersome phrase, "a test with critical function ϕ".) A test is admissible if there is no better test. The decision-theoretic formulation with a two-point action space $A = \{a_0, a_1\}$ and a loss function of the form

$$L(\theta, a_j) = \ell(\theta) > 0 \quad \text{if} \quad \theta \notin \Theta_j, \quad = 0 \quad \text{if} \quad \theta \in \Theta_j \ ,$$

yields the same ordering among tests, and hence the same collection of admissible tests.

Let

(2) $$U_r = U_r(\Theta, \theta_0)$$

$$= \left\{u: ||u|| = 1, \ \exists\ \theta \in \Theta \ni ||\theta|| > r, \text{ and } u = \frac{\theta - \theta_0}{||\theta - \theta_0||}\right\}, \quad r \geq 0 ;$$

and let

(3) $$U(\Theta, \theta_0) = \bigcap_{r \geq 0} U_r(\Theta, \theta_0) \text{ and } U^*(\Theta, \theta_0) = \bigcap_{r \geq 0} \bar{U}_r(\Theta, \theta_0) .$$

Note that if Θ is a closed cone then $U = U^*$; more generally $U \subset U^*$. It is possible that $U = \phi$ but $U^* \neq \phi$.

If $S \subset R^k$ is a convex set let

(4) $$\alpha(u) = \alpha_S(u) = \sup \{x \cdot u: \ x \in S\} .$$

This function is defined for $u \in R^k$,, although we will mainly be interested in its values for $||u|| = 1$. As is well known,

(5) $$\bar{S} = \bigcap_{\{u: ||u||=1\}} \bar{H}^-(u, \alpha_S(u)) .$$

It is clear from the definition (4) that $\alpha(\cdot)$ is lower semicontinuous.

The following lemma is a key result which leads directly to the first main theorem. A result of this type was first proved and used by Birnbaum (1955) in the case of testing for a normal mean. A general result similar to the following lemma was then proved and applied in Stein (1956b).

<u>7.13 Lemma</u>

Fix $\theta_2 \in R^k$. Let

(1) $$S = \bigcap_{u \in U^*} \bar{H}^-(u, \alpha_S(u))$$

where $U^* = U^*(\Theta_1, \theta_2)$. Assume further either that

(2) $$S = \bigcap_{u \in U} \bar{H}^-(u, \alpha_S(u)) , \qquad (U = U(\Theta, \theta_2)) ,$$

or $\alpha_S(u)$ is continuous at u for all $u \in U^* - U$. Let $\phi_1(x) = 1$ for all $x \notin S$. Suppose ϕ_2 is as good as ϕ_1. Then $\phi_2(x) = 1$ for $x \notin S$, a.e. (ν).

(Note: A more formal way to state the conclusion of the lemma is $\nu\{x\colon\ x \notin S,\ \ \phi_2(x) < 1\} = 0$.)

Proof. Assume for convenience $\theta_2 = 0$. Suppose the conclusion of the lemma is false. Then there is an $\varepsilon_0 > 0$, $u_0 \in U^*$ such that

(3) $$C_0 = \{x\colon\ \phi_2(x) \le 1 - \varepsilon_0\}$$

satisfies

(4) $$\nu(C_0 \cap H^+(u_0, \alpha(u_0))) > 0 .$$

Assume $u_0 \in U$. Then there is a sequence $\{\rho_i\}$ with $\rho_i \to \infty$ such that $\{\rho_i u_0\colon\ i=1,\ldots\} \subset \Theta_1$. Theorem 7.7 yields

(5) $$\frac{1 - \pi_{\phi_1}(\rho_i u_0)}{1 - \pi_{\phi_2}(\rho_i u_0)} \le \frac{\int_{\bar{H}^-(u_0,\alpha(u_0))} e^{\theta\cdot x}\,\nu(dx)}{\varepsilon_0 \int_{C_0 \cap H^+(u_0,\alpha(u_0))} e^{\theta\cdot x}\nu(dx)}$$

$$\le C_0 \exp(-\rho_i \varepsilon\ ) \to 0 \quad \text{as} \quad i \to \infty .$$

Hence $\pi_{\phi_1}(\rho_i u_0) > \pi_{\phi_2}(\rho_i u_0)$ for i sufficiently large, which shows that ϕ_2 is not better than ϕ_1.

Now assume $u_0 \notin U$ but $\alpha_S(u)$ is continuous at $u_0 \in U^* - U$. Then $\varepsilon_0 > 0$ in (3) can be chosen small enough so that

(6) $$\nu(C_0 \cap H^+(u, \alpha(u))) > \varepsilon_0$$

for all $||u||=1$ with $||u-u_0|| < \varepsilon_0$. Theorem 7.7, including formula 7.7(4) for the constant c appearing in 7.7(3), now yields, for $\theta = \rho u \in N$,

(7)
$$\frac{1-\pi_{\phi_1}(\rho u)}{1-\pi_{\phi_2}(\rho u)} \le \frac{\int_{\bar{H}^-(u,\alpha(u))} e^{\rho u\cdot x}\,\nu(dx)}{\int_{C_0\cap H^+(u,\alpha(u))} e^{\rho u\cdot x}\,\nu(dx)}$$

$$\le (1/\varepsilon_0)e^{-\rho\varepsilon_0}$$

for $||u|| = 1$ with $||u-u_0|| < \varepsilon_0$. $u_0 \in U^*(\Theta_1)$ implies there exists a sequence $\theta_i \in \Theta_1$ with $||\theta_i|| \to \infty$ such that $\theta_i/(||\theta_i||) \to u_0$. It follows from (7) that $\pi_{\phi_1}(\theta_i) > \pi_{\phi_2}(\theta_i)$ for i sufficiently large. Consequently ϕ_2 is not better than ϕ_1.

It follows from the two cases treated above that ϕ_2 better than ϕ_1 implies $\phi_2(x) = 1$ for (a.e.) $x \notin S$. ||

Lemma 7.13 leads directly to a criterion which can often be used to prove admissibility of conventional tests for appropriate testing problems.

7.14 Corollary

Let $\{p_\theta: \theta \in \Theta\}$, $\Theta = \Theta_0 \cup \Theta_1$ be a standard exponential family, as in 7.12. Let $\theta_2 \in R^k$ and

(1)
$$S = \bigcap_{u\in U^*} \bar{H}^-(u, \alpha_S(u))$$

where $U^* = U^*(\Theta_1, \theta_2)$, as in 7.13(1). Assume (also as in 7.13) that 7.13(2) is satisfied or that $\alpha_S(u)$ is continuous at u for all $u \in U^* - U$. Let $\phi(x) = 1 - \chi_S(x)$ (= 0 if $x \in S$, =1 if $x \notin S$). Then ϕ is an admissible test.

Proof. Suppose ϕ' is any test as good as ϕ. Then, $\phi'(x) = \phi(x) = 1$ for a.e.(ν) $x \notin S$ by Lemma 7.13. But then, $\pi_{\phi'}(\theta_0) \le \pi_\phi(\theta_0)$ implies $\phi'(x) = \phi(x) = 0$ for a.e.(ν) $x \in S$. Thus, $\phi' = \phi$ a.e.(ν). It follows that ϕ is admissible. ||

Remark. It follows from Corollary 7.14 that if Θ_0 is a bounded null hypothesis and $\Theta = R^k$ then any nonrandomized test with convex acceptance region is admissible. When $\Theta_0 = \{\theta_0\}$ is simple and ν is dominated by Lebesgue measure such tests in fact form a minimal complete class -- i.e. a test is admissible if and only if it is nonrandomized and has convex acceptance region (a.e.(ν)). This is the fundamental result which was proved by Birnbaum (1955). See Exercise 7.14.3.

7.15 Application (Univariate general linear model)

Here is a customary canonical form for the normal theory general linear model: $Y \in R^p$ has the normal $N(\mu, \sigma^2 I)$ distribution, $\mu_{s+1} = \ldots = \mu_p = 0$, $\sigma^2 > 0$, and the null hypothesis to be tested is that $\mu_1 = \ldots = \mu_r = 0$, $1 \le r \le s \le p$. (See, e.g. Lehmann (1959, Chapter 7).) This can be reduced via sufficiency and change of variables to a testing question of the form considered above. Let $X_i = Y_i$, $i=1,\ldots,s$, $X_{s+1} = \sum_{j=1}^{p} Y_j^2$. Then the distributions of $X = (X_1,\ldots,X_{s+1})$ form a minimal standard exponential family with canonical parameters $\theta_i = \mu_i/\sigma^2$, $i=1,\ldots,s$, $\theta_{s+1} = -1/2\sigma^2$. The null hypothesis is, therefore, $\Theta_0 = \{\theta \in N: \theta_i = 0, \ i=1,\ldots,r\}$, so that $\Theta_1 = \{\theta \in N: \sum_{i=1}^{r} \theta_i^2 > 0\}$, where of course $N = \{\theta \in R^{s+1}: \ \theta_{s+1} < 0\}$.

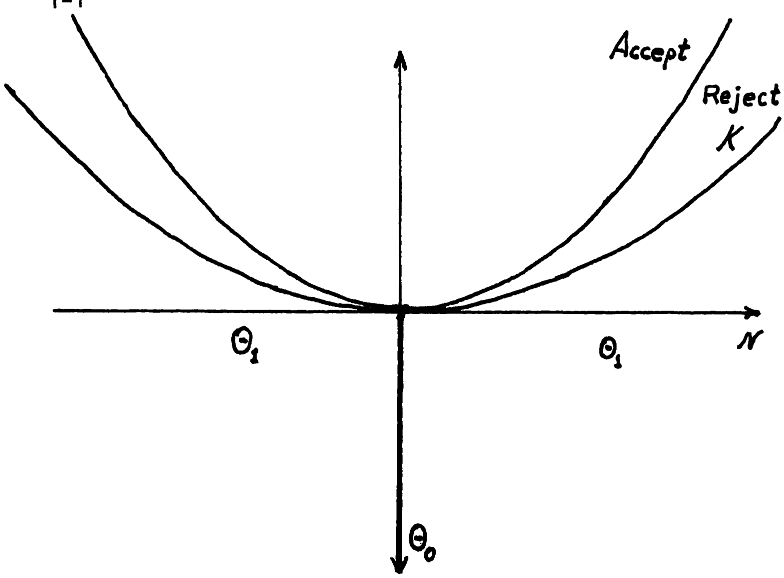


Figure 7.15(1): The F-test when $r = 1 = s$, $p = 2$.

The usual likelihood ratio F-test accepts if (and only if)

(1) $$\frac{\sum_{j=1}^{r} Y_j^2/r}{\sum_{s+1}^{p} Y_j^2/(p - s)} \le F_\alpha ,$$

as determined from tables of the F-distribution. In terms of the canonical variables this region is

(2) $$\frac{\sum_{j=1}^{r} X_j^2}{X_{s+1} - \sum_{j=1}^{r} X_j^2 - \sum_{r+1}^{s} X_j^2} \le \frac{rF_\alpha}{(p - r)} ,$$

or

(3) $$K \sum_{j=1}^{r} X_j^2 + \sum_{r+1}^{s} X_j^2 \le X_{s+1} \quad \text{where} \quad K = 1 + (p - s)/rF_\alpha > 1.$$

(The simple situation for $r = 1 = s$, $p = 2$ is illustrated in Figure 7.15(1), above, which shows K in the upper half-space and N in the lower half. Compare Figures 7.11(1) and Figure 7.12.3.)

Consider a point z in the boundary of the acceptance region (3). Thus, $K \sum_{1}^{r} z_j^2 + \sum_{r+1}^{s} z_j^2 = z_{s+1}$. The outward normal at z is $v = (2Kz_1, \ldots, 2Kz_r, 2z_{r+1}, \ldots, 2z_s, -1)$. Except for the $(s + 1 - r)$ dimensional set having $\sum_1^r z_j^2 = 0$ all positive multiples of this vector lie in Θ_1. It follows that 7.13(1) and 7.13(2) are satisfied (for any choice of $\theta_0 \in \Theta_0$). Thus the F-test (1) (or (2)) is admissible. Note that the test remains admissible by the same reasoning if Θ_1 is restricted by $\sum_{i=1}^{r} \mu_i^2 > a\sigma^2$ since then

$$\Theta_1 = \{\theta \in N: \ \sum_{i=1}^{r} \theta_i^2 > -2\ a\ \theta_{s+1}\}.$$

The same style of reasoning can be used to prove admissibility of a wide variety of tests involving the univariate and multivariate general linear model. It was used in Stein (1956b) to prove admissibility of

Hotelling's T^2 test; Giri (1977) contains a compilation of other results provable by this method, and further references.

7.16 Discussion

If a test is shown to be admissible by virtue of Theorem 7.14 this does not, in itself, constitute a strong recommendation in favor of the test. In principle the following situation may exist: there may be another test ϕ' with $\pi_{\phi'}(\theta) \le \pi_\phi(\theta)$ for all $\theta \in \Theta_0$ and with $\pi_{\phi'}(\theta) \ge \pi_\phi(\theta)$ for "most" $\theta \in \Theta_1$. It might occur that $\pi_{\phi'}(\theta_i) > \pi_\phi(\theta_i)$ for $\theta \in \Theta_1$ except when both $\pi_{\phi'}$ and π_ϕ are very nearly one. In such a case ϕ' would dominate ϕ for all practical purposes.

Of course, a procedure whose admissibility can be proved by Theorem 7.14 may also be a desirable one. The F-test of 7.15 is a good example of this. It is admissible from several perspectives in addition to that of Theorem 7.14. The most surprising of these properties is undoubtedly the fact that it is a Bayes test. See Kiefer and Schwartz (1965) and Exercise 7.16.2.

The F-test is also locally optimal (D-optimality) in the sense that it maximizes (among level-α tests)

$$\min_{\mu\in\Theta_0} \sigma^2 \sum_{i=1}^{r} \frac{\partial^2}{\partial\mu_i^2} \pi_\phi(\mu, \sigma^2) \quad . \tag{1}$$

See Giri and Kiefer (1964) or Giri (1977) and Exercise 7.16.3. When r = s the F-test, ϕ_F, is also optimal in the sense that for any constant $c > 0$ and any level-α test ϕ

$$\min \{\pi_{\phi_F}(\mu, \sigma^2): \sum_{i=1}^{r} \mu_i^2/\sigma^2 = c^2\} \tag{2}$$

$$\ge \min \{\pi_\phi(\mu, \sigma^2): \sum_{i=1}^{r} \mu_i^2/\sigma^2 = c^2\}$$

with equality only if $\phi = \phi_F$. Note that the left side of (2) is a constant. See Brown and Fox (1974b). Brown and Fox (1974a) yields the same result for $s + 1 = r$. For $r \le s + 2$ it is only known that the (minimax) inequality (2) is valid without the (admissiblity) assertion of equality only if $\phi = \phi_F$. This (minimax) assertion follows from the Hunt-Stein theorem as stated in Lehmann (1959).

The next lemma is needed for the complete class theorems which follow it. The lemma can be viewed as an elaboration of Theorem 2.17.

<u>7.17 Lemma</u>

Let ω_n be a sequence of (locally finite) measures concentrated on $\Theta \subset R^k$. Then there exists a subsequence $\omega_{n'}$, a closed convex set S, and a (locally finite) measure ω concentrated on $\bar{\Theta}$ such that

$$(1) \qquad \begin{aligned} \lambda_{\omega_{n'}}(b) &\to \lambda_{\omega}(b)\ , & b \in S^{\circ} \\ \lambda_{\omega_{n'}}(b) &\to \infty \ , & b \notin S\ . \end{aligned}$$

If $\omega_{n'}$, ω, and S are as in (1) and $\theta_2 \in R^k$ then

$$(2) \qquad \bar{S} = \bigcap_{u \in U^*} \bar{H}^-(u, \alpha_S(u)) \quad ,$$

where $U^* = U^*(\Theta, \theta_2)$. (This is similar to 7.13(1).)

Proof. The first part of the lemma is a direct consequence of Theorem 2.17. To prove (2) let $T = \bigcap_{u \in U^*} \bar{H}^-(u, \alpha_S(u))$ and suppose $y \in T^{\circ}$. Then for every $u \in U^*$ there is an $x(u) \in S$ such that $u \cdot x(u) > u \cdot y$.

Define N(u) by

$$(3) \qquad N(u) = \{v: \ ||v|| = 1, \ v \cdot x(u) > v \cdot y\} \quad .$$

N(u) is a relatively open subset of the unit sphere and $u \in N(u)$. Hence $\bigcup_{u \in U^*} N(u) \supset U^*$, and there is a finite subset $u_1, \ldots, u_r \subset U^*$ such that

$$(4) \qquad \underline{\underline{N}} = \bigcup_{i=1}^{r} N(u_i) \supset U^* \quad .$$

For convenience let $x_i = x(u_i)$. Now,

$$(5) \qquad \sup\{||\theta||: \ \theta \in \Theta\ , \ \frac{\theta}{||\theta||} \notin \underline{\underline{N}}\} = B < \infty \quad ;$$

otherwise there would be a sequence $v_i \notin \underline{N}$ with $v_i \to v$ ($v \notin \underline{N}$ since $\underline{N}$ is open) and a sequence $\rho_i \to \infty$ such that $\rho_i v_i \in \Theta$, $i=1,\ldots$; but then $v \in U^* \subset \underline{N}$, a contradiction. Then

$$(6) \qquad \int e^{\theta\cdot y}\, \omega_{n'}(d\theta) \le e^{B||y||}\, \omega_{n'}(\{\theta: ||\theta|| \le B\}) + \sum_{i=1}^{r} \int e^{\theta\cdot x_i}\, \omega_{n'}(d\theta)$$

$$\le e^{B||y||}\, e^{B||x_1||}\, \lambda_{\omega_{n'}}(x_1) + \sum_{i=1}^{r} \lambda_{\omega_{n'}}(x_i)$$

by (3), (4), (5) and the simple fact that

$$\omega_{n'}(\{\theta: ||\theta|| \le B\}) \le e^{B||x_1||} \int e^{\theta\cdot x_1} \omega_{n}'(d\theta) .$$

It follows from (6) and (1) that $y \in S$. Hence $T^\circ \subset \bar{S}$. Since T and $\bar{S}$ are closed and convex this implies $T = \bar{S}$. ||

Here is the complete class theorem from Farrell (1968). It applies to situations where Θ_0 is compact and Θ_0 and Θ_1 are separated sets. See Theorem 7.19 for a partial converse. Results like Theorem 7.18 and 7.19 have been proved in contexts somewhat more general than ordinary exponential families. See Schwartz (1967), Oosterhoff (1969), Ghia (1976), Perlman (1980), and Marden (1982a, 1982b), for such extensions and various applications. In the following statement $\bar{\Theta}_1$ denotes the closure in R^k, not merely the closure relative to N.

<u>7.18 Theorem</u>

Let $\Theta_0 \subset N$ be compact and assume $\Theta_0 \cap \bar{\Theta}_1 = \phi$. Let ϕ' be an admissible test. Then there exists an equivalent test ϕ (i.e. $\pi_{\phi'}(\theta) = \pi_{\phi}(\theta)$, $\theta \in \Theta_0 \cup \Theta_1$), a convex set S satisfying 7.17(2), and a (locally finite) measure H_i on $\bar{\Theta}_i$, $i=0,1$, such that $\lambda_{H_1}(x) < \infty$ for $x \in S^\circ$ and

$$(1) \qquad \phi(x) = \begin{cases} 1 & \text{if } x \notin \bar{S} \\ 1 & \text{if } x \in S^\circ, \quad \dfrac{\lambda_{H_1}(x)}{\lambda_{H_0}(x)} > 1 \\ 0 & \text{if } x \in S^\circ, \quad \dfrac{\lambda_{H_1}(x)}{\lambda_{H_0}(x)} < 1, \end{cases}$$

a.e.(ν). ($\lambda_{H_0}(x)$ is finite since $H_0(\Theta_0) < \infty$.) If $(\Theta_0 \cup \bar{\Theta}_1)^0 \neq \phi$ then $\phi = \phi'$; and hence all admissible tests are of the form ϕ in (1).

Proof. If ϕ' is admissible then according to Theorem 4A.10 there exists an equivalent test ϕ and a sequence of *a priori* distributions G_n (concentrated on finite subsets of Θ) whose Bayes procedures ϕ_n (say), converge to ϕ in the topology of 4A.2. By Exercise 4A.2.1 this convergence means that $\phi_n \to \phi$ weak* -- i.e.

$$(2) \qquad \int (\phi_n(x) - \phi(x))\, g(x)\nu(dx) \to 0$$

for every ν integrable function g. A consequence of (2) is that if a subsequence of $\phi_n(x)$ converges pointwise on some (measurable) subset $T \subset K$ (say $\phi_{n'}(x) \to \lambda(x)$, $x \in T$) then the limit must be ϕ (i.e., $\phi(x) = \lambda(x)$, $x \in T$, a.e.(ν)).

Let

$$(3) \qquad H_{in}(d\theta) = e^{-\psi(\theta)}\, G_n(d\theta) \Big/ \int_{\Theta_0} e^{-\psi(\theta)}\, G_n(d\theta), \qquad \theta \in \Theta_i, \qquad i = 0, 1 .$$

Note that $H_{0n}(\Theta_0) = 1$. Then

$$(4) \qquad \phi_n(x) = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \quad \text{if} \quad \frac{\lambda_{H_{1n}}(x)}{\lambda_{H_{0n}}(x)} \left\{ \begin{matrix} > & 1 \\ < & 1 \end{matrix} \right. .$$

Let $\omega_n = H_{0n} + H_{1n}$. Let $\omega_{n'}$, ω, S be as in Lemma 7.17. Let $H_i = \omega_{|\Theta_i}$, i=0, 1, so that $H_{in'} \to H_i$, i=0, 1, as $n' \to \infty$. Then $H_0(\Theta) = H_0(\Theta_0) = 1$ since

$H_{0n}(\Theta_0) = 1$ and Θ_0 is compact. The assertions in (1) follow from this along with Lemma 7.17, (4), and the decision theoretic facts in the first paragraph of the proof.

If ϕ' and ϕ are equivalent and $(\Theta_0 \cup \bar{\Theta}_1)^\circ \neq \phi$ then $\phi' = \phi$ (a.e.(ν)) by completeness (Theorem 2.12). ||

Many of the tests produced by the recipe 7.18(1) are admissible. In certain statistical situations, it can even be concluded that all of them are admissible. Then Theorem 7.18 describes the minimal complete class. The following converse to Theorem 7.18 contains statements of these facts. It is not entirely satisfactory but it is the best general result we have been able to devise. For the purpose of this theorem define

$$(1) \qquad \Theta_1^* = \{\theta_1 \in \bar{\Theta}_1 : \ \theta_1 \in N \quad \text{or there is a} \quad \theta_2 \in \Theta_1 \ni (1-\rho)\theta_2 + \rho\theta_1 \in \Theta_1 \quad \text{for} \quad 0 \leq \rho < 1\} \quad .$$

(See Exercise 7.19.3 for an extension of (1).)

7.19 Theorem

Consider the testing problem described in Theorem 7.18.

Suppose ϕ satisfies 7.18(1) where H_1 is concentrated on Θ^* and $\bar{S}$ satisfies all the assumptions of Lemma 7.13 relative to some $\theta_2 \in R^k$. Suppose also that

$$(2) \qquad \phi(x) = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \text{ if } \quad x \in \bar{S} \quad \text{and} \quad \frac{\lambda_{H_1}(x)}{\lambda_{H_0}(x)} \begin{cases} > 1 \\ < 1 , \end{cases} \quad \text{a.e.}(\nu) \quad .$$

(This is a mild extension of the latter part of 7.18(1).) Then any critical function as good as ϕ must also satisfy (2) and 7.18(1) with the same values of S, H_0, H_1. If also either

$$(3') \qquad \nu(\{x: \frac{\lambda_{H_1}(x)}{\lambda_{H_0}(x)} = 1 \quad \text{and} \quad \phi(x) < 1\}) = 0 \quad ,$$

or

$$(3'') \qquad \nu(\{x: \frac{\lambda_{H_1}(x)}{\lambda_{H_2}(x)} = 1 \text{ and } \phi(x) > 0\}) = 0 \quad ,$$

or

$$(4) \qquad (\text{Supp }(H_0 + H_1))^\circ \neq \phi$$

then ϕ is admissible; and if η is as good as ϕ then $\eta = \phi$ a.e.(ν).

If ν is dominated by Lebesgue measure, $U(\Theta, \theta_2) = U^*(\Theta, \theta_2)$ for some $\theta_2 \in R^k$, and $\bar{\Theta}_1 \subset \Theta_1^*$ then the collection of tests of the form 7.18(1) is a minimal complete class.

Proof. Suppose ϕ is defined by (2) and 7.18(1) where $\bar{S}$ satisfies the assumptions of Lemma 7.13. Suppose η is another critical function as good as ϕ. Then $\eta(x) = 1$ if $x \notin \bar{S}$ by Lemma 7.13.

If $\theta \in \Theta_1$ then

$$(5) \qquad 0 \leq \int(\eta(x) - \phi(x))e^{\theta\cdot x}\,\nu(dx)$$

since $0 \leq \pi_\eta(\theta) - \pi_\phi(\theta)$. By continuity (5) also holds if $\theta \in \bar{\Theta}_1 \cap N$. Now, suppose $\zeta_\rho = (1 - \rho)\theta_2 + \rho\theta_1 \in \Theta_1$ for $0 \leq \rho < 1$. Then (5) holds at $\theta = \zeta_\rho$ and $\int(\eta(x) - \phi(x))e^{\zeta_\rho\cdot x}\,\nu(dx)$ is continuous in ρ as $\rho \uparrow 1$ by Exercise 1.13.1(ii). It follows that (5) holds whenever $\theta \in \Theta_1^*$.

The opposite inequality to (5) holds when $\theta \in \Theta_0$, and H_0 is finite since $\bar{\Theta}_0 \subset N$ is compact. Hence

$$(6) \qquad 0 \leq \int (\int(\eta(x) - \phi(x))e^{\theta\cdot x}\,\nu(dx))(H_1(d\theta) - H_0(\theta)) \quad .$$

Notice that $\eta(x) - \phi(x) \leq 0$ whenever $\lambda_{H_1}(x) > \lambda_{H_0}(x)$, so that

$$\int (\eta(x) - \phi(x))^+ \lambda_{H_1}(x)\,\nu(dx) \leq \int(\eta(x) - \phi(x))^+ \lambda_{H_0}(x)\,\nu(dx)$$

$$= \int\int e^{\theta\cdot x}\,\nu(dx)\,H_0(d\theta) < \infty \quad .$$

Furthermore, as already noted, H_0 is a finite measure. Hence the order of integration in (6) can be reversed, yielding that

$$(7) \qquad 0 \leq \int_{\bar{S}} (\eta(x) - \phi(x))(\lambda_{H_1}(x) - \lambda_{H_0}(x))\nu(dx) < \infty \quad ;$$

with the integral extending only over the region $x \in \bar{S}$ since $\eta(x) = \phi(x)$ for $x \notin \bar{S}$. Because ϕ satisfies (2), the integrand in (7) is non-positive; hence $\eta(x)$ also satisfies (2), for otherwise the integral would be negative.

If in addition ϕ satisfies (3') then $\pi_\phi(\theta_1) > \pi_\eta(\theta_1)$, $\theta_1 \in \Theta_1$, (a contradiction) unless $\eta(x) = \phi(x)$ a.e. (ν). Similarly if (3") is satisfied $\eta(x) = \phi(x)$ a.e.(ν); for otherwise $\pi_\phi(\theta_0) < \pi_\eta(\theta_0)$, $\theta_0 \in \Theta_0$. Finally, suppose (4) is satisfied in place of (3') or (3"). Note that the reasoning following (7) shows that equality holds in (7) and hence in (6). From this it follows that $\int(\eta(x) - \phi(x))e^{\theta\cdot x}\nu(dx) = 0$ a.e. $H_0 + H_1$ since this integral is non-negative on Θ_1^* and non-positive on Θ_0. (4) then implies $\eta(x) = \phi(x)$ a.e. (ν) by completeness and hence ϕ is admissible. This completes the proof of all assertions in the middle paragraph of the theorem.

If ν is dominated by Lebesgue measure and also satisfies the remaining assumptions of the last paragraph of the theorem then

$$\nu\{x: \frac{\lambda_{H_1}(x)}{\lambda_{H_0}(x)} = 1\} = 0$$

so that any test, ϕ, of the form 7.18(1) is also of the form (2), and (3') (and (3")) is satisfied, and H_1 is concentrated on $\bar{\Theta}_1 \subset \Theta_1^*$, and S satisfies assumption 7.13(2) of Lemma 7.13. It follows that ϕ is admissible. ||

COMPLETE CLASS THEOREMS FOR TESTS (Contiguous Hypotheses)

7.20 Definitions:

It is necessary to characterize the local structure of Θ_1 near Θ_0.

Let $\Theta_0 = \{\theta_0\}$ and Θ_1 be given and let

(1) $$\mathcal{J}(\varepsilon) = \{J\colon J \text{ is a finite non-negative measure on } \{\theta\colon \theta \in \Theta_1, \|\theta-\theta_0\| \le \varepsilon\},\ \int J(d\theta) \le 1,\ \int \frac{J(d\theta)}{\|\theta-\theta_0\|^2} < \infty,\ \left\| \int \frac{\theta-\theta_0}{\|\theta-\theta_0\|^2} J(d\theta) \right\| \le 1\}$$

Then let

(2) $$\Delta(\varepsilon) = \{(v,M)\colon v = \int \frac{\theta-\theta_0}{\|\theta-\theta_0\|^2} J(d\theta),\quad M = \int \frac{(\theta-\theta_0)(\theta-\theta_0)'}{\|\theta-\theta_0\|^2} J(d\theta),\ J \in \mathcal{J}(\varepsilon)\}.$$

Also, let $\Delta = \bigcap_{\varepsilon>0} \bar{\Delta}(\varepsilon)$. Note that $v \in R^k$ and M is a positive semidefinite $k \times k$ matrix, and Δ and $\bar{\Delta}(\varepsilon)$ are compact, convex sets.

In various typical statistical problems it is not hard to explicitly describe Δ. For example, if $\theta_0 = 0$ and $\bar{\Theta} = \theta_0 \cup \bar{\theta}_1$ is a closed conical set then Δ is the convex hull of points of the form

(3) $$\begin{aligned} &(v,0)\colon v \in \bar{\Theta},\ \|v\| \le 1\ , \text{ and} \\ &(0,M)\colon M = vv' \ni v \in \bar{\Theta},\ -v \in \bar{\Theta},\ \|v\| \le 1\ . \end{aligned}$$

(See Exercise 7.20.1.) As another example, suppose Θ is a twice differentiable curved exponential family at Θ_0. This means that there are two orthogonal vectors u_1, $u_2 \in R^k$, with $\|u_1\| = 1$ such that for $\theta \in \Theta$

(4) $$\theta-\theta_0 = ((\theta-\theta_0)\cdot u_1)u_1 + \|\theta-\theta_0\|^2 u_2 + o(\|\theta-\theta_0\|^2).$$

(Note in (4) that $|(\theta-\theta_0)\cdot u_1| = \|\theta-\theta_0\| + o(\|\theta-\theta_0\|^2)$, and also that $u_2 = 0$ is a possible value of u_2.) Then Δ is the convex hull of $(u_1,0)$, $(-u_1,0)$ and (u_2,u_1u_1'). (See Exercise 7.20.2.)

As with earlier results the full complete class characterization is not directly as a generalized Bayes test but involves an extension of this notion. As part of this extension the kernel $e^{\theta\cdot x}$ is replaced by

(5) $$\omega(\theta,x) = \frac{e^{\theta\cdot x}-1-\theta\cdot x}{||\theta||^2} .$$

A converse result which sometimes yields a characterization of the minimal complete class is given in Theorem 7.22. As with earlier results both of the following theorems can be profitably extended beyond the exponential family context in which they are proved below. See Marden and Perlman (1980), Marden (1981, 1982b), Cohen and Marden (1985), Brown and Sackrowitz (1984, Theorem 6.1), and Brown and Marden (in preparation).

7.21 Theorem

Let $\Theta_0 = \{\theta_0\}$ be a simple null hypothesis. Let ϕ' be an admissible test of Θ_0 versus Θ_1. Then there exists an equivalent test ϕ and a closed convex set S satisfying 7.17 (2) such that

(1) $$\phi(x) = 1 \quad x \notin S.$$

Further, for every $x_0 \in S^\circ$ there is a finite non-negative measure H on $\bar{\Theta}_1 - \{\theta_0\}$ with $S^\circ \subset \{x: \int e^{\theta\cdot(x-x_0)}H(d\theta) < \infty\}$, a constant $C \in R$, an $M \in \Delta_2$, and a $v \in R^k$ satisfying (3), below, with at least one of C, H, v, M being non-zero, such that for all $x \in S^\circ$

(2) $$\phi(x) = \begin{cases} 1 & \text{if } C \;\begin{matrix}<\\ \\ >\end{matrix}\; \int \omega(\theta,x-x_0)e^{\theta\cdot x_0}H(d\theta)+v\cdot(x-x_0) \\ & \qquad + (x-x_0)'M(x-x_0)/2 . \\ 0 & \end{cases}$$

If $\Theta_1^\circ \neq \phi$ then $\phi = \phi'$ a.e. (ν).

Define

$$v_\varepsilon = v - \int_{||\theta-\theta_0||>\varepsilon} \frac{\theta-\theta_0}{||\theta-\theta_0||^2} H(d\theta).$$

Then there is a sequence $\varepsilon_j \to 0$ such that $\lim_{j\to\infty} v_{\varepsilon_j} = v_0$ (say) exists, and

(3) $$(v_0,M) \in \Delta.$$

(Note that if $\int||\theta||^{-1}H(d\theta) < \infty$ the extreme right side of (2) can be rewritten as

$$(3'') \qquad \int \frac{e^{\theta\cdot(x-x_0)}-1}{||\theta||^2} H(d\theta) + v_0\cdot(x-x_0)'M(x-x_0)/2 \ .$$

In particular, $\lim_{\varepsilon\to 0} v_\varepsilon = v_0$ exists.)

Proof. The assertion just after (2) follows from completeness, as in Theorem 7.17. Now, suppose ϕ' is admissible. Then by Theorem 4A.10 there is an equivalent ϕ and a sequence of prior distributions G_i concentrated on finite subsets of Θ such that the Bayes procedures, $\phi_i = \phi_{G_i}$, converge to ϕ in the topology of 4A.2. (See the proof of Theorem 7.18 for further remarks.)

Without loss of generality let $\theta_0 = 0$ and

$$G_i'(d\theta) = \frac{e^{-\psi(\theta)}G_i(d\theta)}{e^{-\psi(0)}G_i\{0\}} \ .$$

Thus $\phi_i(x) = \{{1 \atop 0}\}$ according to whether

$$(4) \qquad \int_{\Theta_1} e^{\theta\cdot x}G_i'(d\theta) \ {> \atop <} \ 1 \ .$$

As in 7.17, it is possible to reduce $\{G_i'\}$ to a subsequence (if necessary) such that now for some closed S satisfying 7.17 (2),

$$\lim \int e^{\theta\cdot x}G_i'(d\theta) = \infty \qquad x \notin S$$

$$0 \le \lim \int e^{\theta\cdot x}G_i'(d\theta) = q(x) < \infty \qquad x \in S^\circ \ ,$$

where $G_i' \to G'$ and $q(x) = \int e^{\theta\cdot x}G'(d\theta)$. Clearly, (1) is satisfied.

Assume without loss of generality that $x_0 = 0 \in S^\circ$. Rewrite (4) as

$$(5) \qquad \int_{\Theta_1} \omega(\theta,x)||\theta||^2G_i'(d\theta) + \int_{\Theta_1} (\theta\cdot x)G_i'(d\theta) \ {> \atop <} \ c_i' \ .$$

Let $d_i = \int||\theta||^2G_i'(d\theta) + ||\int \theta G_i'(d\theta)|| + |c_i'|$ and $H_i(d\theta) = d_i^{-1}||\theta||^2G_i'(d\theta)$. Substituting in (5) and multiplying through by d_i^{-1} yields

$$(6)\qquad \int_{\Theta_1} \omega(\theta,x)H_i(d\theta) + \int_{\Theta_1} \frac{\theta\cdot x}{\|\theta\|^2} H_i(d\theta) \gtrless c_i$$

where $\int H_i(d\theta) + \|\int(\theta/\|\theta\|^2)H_i(d\theta)\| + |c_i| = 1$. Reduce $\{H_i\}$ to a subsequence (if necessary) so that

$$(7)\qquad \int(\theta/\|\theta\|^2)H_i(d\theta) \to v, \quad c_i \to C.$$

$H_i \to H'$ since $G_i' \to G'$. Furthermore $\int H'(d\theta) + ||v|| + C = 1$ since $x_0 = 0 \in S^{\circ}$. Let $H = H'|_{\overline{\Theta}_1 - \{0\}}$.

Let $\varepsilon > 0$ such that $H(\{\theta: \|\theta\| = \varepsilon\}) = 0$. (All but a countable set of ε's satisfy this.) For each $x \in S^0$

$$(8)\qquad \int_{\|\theta\|>\varepsilon} \omega(\theta,x)H_i(d\theta) \to \int_{\|\theta\|>\varepsilon} \omega(\theta,x)H(d\theta),$$

and

$$(9)\qquad \left|\int_{\|\theta\|\le\varepsilon} \left(\omega(\theta,x) - \frac{x'\theta\theta'x}{2\|\theta\|^2}\right)H_i(d\theta)\right| = O(\varepsilon)$$

since $(e^t - 1 - t - t^2/2)/t^2 = O(t)$ and $\int H_i(d\theta) \le 1$. Another subsequence may now be taken, if necessary, so that the following limits exist:

$$(10)\qquad v_\varepsilon = \lim_{i\to\infty} \int_{\|\theta\|\le\varepsilon} \frac{\theta}{\|\theta\|^2} H_i(d\theta) = v - \int_{\|\theta\|>\varepsilon} \frac{\theta}{\|\theta\|^2} H(d\theta)$$

$$(11)\qquad M_\varepsilon = \lim_{i\to\infty} \int \frac{\theta\theta'}{\|\theta\|^2} H_i(d\theta).$$

By definition, $(v_\varepsilon, M_\varepsilon) \in \overline{\Delta}(\varepsilon)$. $\overline{\Delta}(\varepsilon)$ is compact in the obvious topology. Hence there is a subsequence $\varepsilon_j \to 0$ so that $(v_{\varepsilon_j}, M_{\varepsilon_j}) \to (v_0, M) \in \Delta$. If necessary another subsequence of $\{H_i\}$ may be extracted using a diagonalization argument so that (10) and (11) hold for each ε_j. It follows from (5), (7), (8), (9), and (11) that for $x \in S^0$

$$\phi_i(x) \to \begin{matrix}1\\0\end{matrix} \text{ if } C \begin{matrix}<\\>\end{matrix} \int\omega(\theta,x)H(d\theta) + v\cdot x + x'Mx/2 .$$

Note that Tr $M = H'(\{0\})$. Hence Tr $M + H(\bar{\Phi}_1) + ||v|| + C = 1$ so that at least one of M, H, $||v||$, C are non-zero.

It follows from (10) that (3) is satisfied. Since $\phi_i \to \phi$ in the topology of 4A.2 this yields (2). $||$

7.22 Theorem

Consider the testing problem described in Theorem 7.21. Suppose $\theta_0 \in N^\circ$ and ϕ satisfies 7.21(1), (2), and (3) where $\bar{S}$ satisfies all the assumptions of Lemma 7.13 and H is concentrated on Θ_1^*, as defined in 7.19(1). Suppose $\phi(x)$ is also given by 7.21(2) for $x \in \bar{S} - S^\circ$. Then any critical function as good as ϕ must also satisfy 7.21(1) for $x \notin \bar{S}$ and 7.21(2) for $x \in \bar{S}$ (a.e. (ν)) with the same values of S, H, v, M, C, x_0.

If also either

(1) $\nu\{x: \omega(\theta,x-x_0)H(d\theta) + v'\cdot(x-x_0) + (x-x_0)'M(x-x_0)/2 = C,\ \phi(x) < 1\} = 0$

or

(1') $\nu\{x: \omega(\theta,x-x_0)H(d\theta) + v'\cdot(x-x_0) + (x-x_0)'M(x-x_0)/2 = C,\ \phi(x) > 0\} = 0$

or

(2) $$(\text{Supp } H)^\circ \neq \phi$$

then ϕ is admissible; and if η is as good as ϕ then $\eta = \phi$ a.e. (ν).

If ν is dominated by Lebesgue measure, $U(\theta,\theta_2) = U^*(\theta,\theta_2)$ for some $\theta_2 \in R^k$, and $\bar{\Theta}_1 = \Theta_1^*$ then the collection of tests of the form 7.21(1), (2) is a minimal complete class.

Proof. Much of the proof resembles that of Theorem 7.19 (as does much of the statement of the theorem). Assume with no loss of generality that $\theta_0 = 0$ and $x_0 = 0$. Let $\varepsilon_j \to 0$ and v_{ε_j} be as in 7.21(3) and let $J_j \in \mathcal{J}(\varepsilon_j)$, be measures supported on finite subsets such that

(3) $$v_{\varepsilon_j} - \int \frac{\theta}{||\theta||^2} J_j(d\theta) \to 0$$
$$\int \frac{\theta\theta'}{||\theta||^2} J_j(d\theta) \to M.$$

Let $H_{1i} = H|_{\{\theta:\|\theta\|>\varepsilon_i\}} + J_i$, $H_{0i}(\{0\}) = C + \int\|\theta\|^{-2}J_i(d\theta)$. As in 7.19 if η is better than ϕ then η satisfies 7.21 (1) and

(4) $$0 \leqq \int_S (\eta(x)-\phi(x)) \int e^{\theta\cdot x}(H_{1i}(d\theta) - H_{0i}(d\theta))\nu(dx).$$

For each $x \varepsilon S$

(5) $$\begin{aligned}&\int e^{\theta\cdot x}(H_{1i}(d\theta) - H_{0i}(d\theta)) \\ &\quad = \int_{\|\theta\|>\varepsilon_i} \omega(\theta\cdot x)H(d\theta) + v\cdot x + x'Mx/2 - C \\ &\qquad + \left(\int \frac{\theta}{\|\theta\|^2} J_i(d\theta) - v_{\varepsilon_i}\right)\cdot x \\ &\qquad + \left(\int\omega(\theta\cdot x)J_i(d\theta) - x'Mx/2\right).\end{aligned}$$

Lemma 2.1 implies that the dominated convergence theorem can be invoked in (4), (5) as $i \to \infty$ since $0 \varepsilon N^{\circ}$ and $\omega(\theta\cdot x) = O(e^{\theta\cdot x} + 1)$. Hence

(6) $$0 \leqq \int_S (\eta(x)-\phi(x))\left(\int\omega(\theta\cdot x)H(d\theta)+v\cdot x+x'Mx/2-C\right)\nu(dx).$$

It follows that η satisfies 7.21 (2). The remaining assertions of the theorem are proved just as the analogous assertions in Theorem 7.19. ||

EXERCISES

7.2.1

Prove proposition 7.2. [If ν is a finite measure and $\nu(\{||x|| > \alpha\}) = O(e^{-\varepsilon\alpha})$ then $E_\nu(e^{\varepsilon'||x||}) < \infty$ for all $0 < \varepsilon' < \varepsilon$.]

7.4.1

(i) Let S be a convex set with $\rho = \inf\{||x||: x \notin S\}$. Suppose for some $\varepsilon > 0$, $c < \infty$ 7.4(1) holds i.e.

(1) $$P_{\theta_0}(\{X/\alpha \notin S\}) < \quad c\exp(-\varepsilon\rho\alpha) \qquad \forall \quad \alpha \in R \ .$$

Show that $\{\theta: ||\theta - \theta_0|| < \varepsilon'\} \quad N^\circ$ for all $\varepsilon' < \varepsilon$. (ii) Give an example of a nonconvex set with $\nu\{x: ||x|| < \rho, \; x \notin S\} > 0$ and in which (1) holds but $\{\theta: ||\theta - \theta_0|| < \varepsilon'\} \not\subset N^\circ$ for any $\varepsilon' < \varepsilon$.

7.5.1

Let $\theta_0 \in N^\circ$ and $H^+ = H^+(v, \alpha)$. Show

(1) $$\lim_{n\to\infty} (n^{-1} \log P_{\theta_0}(\bar{X}_n \in \bar{H}^+)) \leq -\tilde{K}(\bar{H}^+, \xi(\theta_0)) \quad .$$

[Use Theorem 7.5 and Proposition 5.15.]

7.5.2

In 7.5.1 suppose $\bar{H}^+ \cap K^\circ \neq \phi$. Show

(1) $$\lim_{n\to\infty} (n^{-1} \log P_{\theta_0}(\bar{X}_n \in H^+)) \;=\; -\tilde{K}(H^+, \;\xi(\theta_0)) \quad .$$

[For one direction use 7.5.1(1). For the other let $P_{\tilde{\theta}}^{(n)}$ denote the distribution of S_n under $\tilde{\theta} = \theta(\tilde{\xi}_{H^+}(\theta))$.]

(2) $$P_{\theta_0}(\bar{X} \in H_0) \;\geq\; \exp[-n(K(\tilde{\theta}, \theta_0) + \varepsilon)]\; P_{\tilde{\theta}}^{(n)}(\{S: |(\theta_0 - \tilde{\theta}) \cdot \tfrac{S}{n}| < \varepsilon\})$$

$$\to \; \exp[-n(K(\tilde{\theta}, \theta_0) + \varepsilon)]$$

by the Central Limit Theorem (Exercise 5.15.1).]

7.5.3

Let $\theta_0 = \theta(\xi_0) \in N°$. Let Q be a closed subset of R^k. Show

$$\varlimsup_{n\to\infty} (n^{-1} \log P_{\theta_0}(\bar{X}_n \in Q)) \leq -\tilde{K}(Q, \xi_0) .$$

[Let $\varepsilon > 0$. Show $Q \subset \sum_{i=1}^{I} H^+(v_i, \alpha_i)$ where $\tilde{K}(H^+(v_i, \alpha_i)) \geq \tilde{K}(Q, \xi_0) - \varepsilon$. When $k \geq 2$ this requires some care.) Apply 7.5.2.]

7.5.4

Let $\theta_0 = \theta(\xi_0)$. Let $Q \subset R^k$ be a set such that $\tilde{K}(Q°, \xi_0) = \tilde{K}(\bar{Q}, \xi_0) = \tilde{k}$ (say). Then

$$(1) \qquad \lim_{n\to\infty} n^{-1} \log P_{\theta_0}(\bar{X}_n \in Q) = -\tilde{k} .$$

[Reason as in 7.5.2 and use 7.5.3.]

7.5.5

Let $X_1, \ldots$ be i.i.d. random variables on R^k with distribution F. Let $h: X \to R^k$ be measurable and $Q \subset R^k$. Let $\xi(Q) = \inf \{\xi_F(x): x \in Q\}$ where $\xi_F(x)$ denotes the entropy as defined in 6.16(1). Suppose $\xi(Q°) = \xi(\bar{Q})$ and $E(\exp(\varepsilon||X||)) < \infty$ for some $\varepsilon > 0$. Then

$$\lim_{n\to\infty} n^{-1} \log P(\bar{X}_n \in Q) = E(Q) .$$

7.5.6

(i) Show that $\tilde{K}(\cdot, \xi_0)$ is relatively continuous on $\{x: \tilde{K}(x, \xi_0) < \infty\}$ if $\nu(K - K°) = 0$, if $k = 1$, or if ν is concentrated on a countable number of points satisfying Assumptions in Theorem 6.23. If so, then for Q an open set $\tilde{K}(Q, \xi_0) = \tilde{K}(\bar{Q}, \xi_0)$ as required in 7.5.4. (ii) Given an example where Q is open and $\tilde{K}(Q, \xi_0) \neq \tilde{K}(\bar{Q}, \xi_0)$. [Let ν be Lebesgue measure on the first quadrant of R^2 plus a unit mass at the origin.]

7.7.1

Hwang (1983) raises the following question: Let $X \sim N(\theta, I)$, $\theta \in R^k$. Does there exist an estimator $\delta: R^k \to R^k$ for which

(1) $\quad P_\theta(||\delta(X) - \theta|| \le B) \ge P_\theta(||X - \theta|| \le B) \quad \forall\ B > 0,\ \theta \in R^k$,

with strict inequality for some B, θ? (If so, δ would be said to "stochastically dominate" $\delta_0(x) = x$. Note that for fixed $B > 0$ there exists an estimator δ dominating δ_0 in the sense of satisfying (1) for all $\theta \in R^k$. See Hwang (*op. cit.*) and references cited therein.) It can be shown that $\delta \ne \delta_0$ exists satisfying (1) if and only if there exists a continuous spherically symmetric function $\delta \ne \delta_0$ satisfying (1). Show that no such function exists. [Suppose $||\delta(x_0)|| < ||x_0||$ for some $x_0 \in R^k$ (and hence for a neighborhood of x_0). Let $\theta_\rho = \rho x_0$ and $B_\rho = (\rho - 1)||x_0||$. Show that for some $\varepsilon > 0$, sufficiently small,

(1)
$$\frac{P_{\theta_\rho}(||\delta(X) - \theta_\rho|| > B_\rho)}{P_{\theta_\rho}(||X - \theta_\rho|| > B_\rho)}$$

$$\ge \frac{P_{\theta_\rho}(X \in H^+(x_0, ||x_0||^2),\ \delta(X) \in H^-(x_0, ||x_0||^2))}{e^{\varepsilon\rho}\, P_{\theta_\rho}(X \in H^-(x_0, ||x_0||^2))} \to \infty .$$

Use the multivariate generalization of 7.3(3) to estimate the denominator on the left of (1); then use 7.7(3) for the asymptotic assertion in (1). A similar argument, with different θ_ρ and B_ρ, applies when $||\delta(x_0)|| > ||x_0||$ for some $x_0 \in R^k$. See Brown and Hwang (in preparation).]

7.9.1

Consider the estimation problem described in Exercise 4.24.3. Show that the estimtor 4.24.3(1) is admissible. [Use Theorem 7.7 and Corollary 7.9 to show that if δ' is better than δ then $\delta'(x) = 0$, $x < 1$, and $\delta'(x) \le \frac{1}{2}$, $x = 1$, and symmetrically for $x \ge 2$. Among all such estimators δ minimizes the risk at $\theta = 0$.]

7.9.2 (A uniform version of Corollary 7.9)

Let $V_1 \subset V_2$ be subsets of the unit sphere in R^k with V_1 closed and V_2 relatively open in the unit sphere. Let

$$\alpha(v) = \sup \{\alpha: \; K \cap H^+(v, \alpha) \neq \phi\} \quad . \tag{1}$$

Assume $\alpha(v) < \infty \;\; \forall \; v \in V_2$. Then

(i) $\alpha(v)$ is continuous for $v \in V_2$

(ii) $\forall \; \varepsilon > 0 \;\; \exists \;\; \delta > 0 \;\; \ni \;\; \nu(H^+(v, \alpha(v) - \varepsilon)) > \delta \; , \;\; v \in V_1$

(iii) $\forall \; \varepsilon > 0 \;\; \exists \;\; r_0 \;\; \ni$

$$v \cdot \xi(rv) \; > \; \alpha(v) - \varepsilon \;\; \forall \;\; v \in V_1, \quad r > r_0 \quad . \tag{2}$$

7.9.3

Consider a steep exponential family. Let $K \subset \{x: \; x_1 \leq 0\}$, $0 \in K$, and let K be strictly convex. Let $y \in \partial K$, $y \neq 0$. Let $\theta_i \in N°$, $i=1,\ldots,$ such that $\xi(\theta_i) \to y$. Then, (i) $\exists \; I < \infty$, $\varepsilon > 0$, $\delta > 0$ such that $\nu(H^+(\frac{\theta_i}{||\theta_i||},\varepsilon)) > \delta$ for all $i > I$. Hence, (ii) $\psi(\theta_i) \geq \varepsilon||\theta|| + \ln \delta$ for all $i > I$, and (iii) $\lim_{i\to\infty} \psi(\theta_i) = \infty$.

[There exist V_1, V_2 as in 7.9.2 and $\varepsilon > 0$, $\delta > 0$, satisfying $\alpha(v) < \varepsilon$, $v \in V_2$; $v \cdot y < -2\varepsilon$, $v \in V_1$; and

$$\nu(H^+(v, \varepsilon)) \; > \; \delta \quad \forall \quad v \notin V_1 \quad . \tag{1}$$

(Draw pictures in R^2 to help see why the above is true. The strict convexity is important here.) Now, $||\theta_i|| \to \infty$. (Why?) Hence, $\frac{\theta_i}{|| \theta_i||} \notin V_1$ for i sufficiently large, by 7.9.2(2).]

7.9.4

Consider a steep exponential family. Let $\Theta \subset N$ be relatively closed in N and assume K is strictly convex. Suppose $x \in \partial K$ but $x \notin (\xi(\Theta \cap N°))^-$. Show that $\hat{\theta}(x) \neq \phi$. (This result complements Theorem 5.7. I believe it should be possible to prove it by showing the above hypotheses imply that 5.7(1) is satisfied. However, the hint below indicates a different argument.

[Assume $x = 0 \in K \subset \{x: \; x_1 \leq 0\}$ (w.l.o.g.). Apply 7.9.3 to show $\lim_{||\theta||\to\infty, \theta\in\Theta} \psi(\theta) = \infty$. Now proceed as in the proof of Theorem 5.7, following

5.7(2).]

7.9.5

Consider a standard exponential family with natural parameter space N. Let $v \in R^k$ and $\alpha_0 = \sup\{\alpha: \nu(H^+(v, \alpha)) > 0\}$. Let $\theta_i = \rho_i v + \eta_i$ as in Corollary 7.9. Then

(1) $$\lim_{i\to\infty} v \cdot \nabla\psi(\theta_i) = \alpha_0 .$$

Hence, there exist a $c > -\infty$ such that

(2) $$\psi(\theta_i) \geq -c + \alpha\rho_i ,$$

and, consequently,

(3) $$p_{\theta_i}(x) \to 0 \quad \forall \; x \in H^-(v, \alpha_0) .$$

[The key assertion, (1), is a uniform version of Theorem 3.9, since for $\eta_i \equiv \eta$ it follows immediately from that theorem. However, it seems easier to prove (1) as a consequence of Corollary 7.9. (Alternatively, one may also derive the above, as well as 7.9, through an application of convex duality, since $K° = R$, etc.)]

7.11.1

In the situation in Corollary 7.11 let $\rho(\theta_i) = P_{\theta_i}(S_2)/(P_{\theta_i}(S_1))$. Construct examples (i) in which $\rho(\theta_i) \sim ||\theta_i||^{-\alpha}$, $\alpha > 0$; (ii) in which $\rho(\theta_i) \to 0$ but $||\theta_i||^{\alpha}\rho(\theta_i) \to \infty$ for all $\alpha > 0$; and (iii) in which $\rho(\theta_i) = O(||\theta_i||^{-\alpha})$ for all $\alpha > 0$ but $e^{-\alpha||\theta_i||}\rho(\theta_i) \to \infty$ for all $\alpha > 0$. [(i) Let $k = 1$, $\nu(\{0\}) = 1$ and $\nu(dx) = x^{\alpha-1}\,dx$ on $x > 0$.]

7.12.1

Consider a testing problem, as in 7.12 with $\Theta_0 = H(v, \alpha) \cap N$, $\Theta_1 = N - \Theta_0$, and $\Theta_0 \cap N° \neq \phi$. For $z \in R^k$, let $z = z^{(1)} + z^{(2)}$ where $z^{(1)} \in H(v, \alpha)$, $z^{(2)} = \rho v \perp H(v, \alpha)$. Assume (w.l.o.g.) $\nu(R^k) = 1$. Show

(i) If ϕ' is better than ϕ then

$$\int_{x^{(1)}=y} \phi(x)\nu(dx|x^{(1)} = y) = \int_{x^{(1)}=y} \phi'(x)\nu(dx|x^{(1)} = y) \tag{1}$$

$$y \in H(\nu, \alpha) \quad \text{a.e.}(\nu)$$

and

$$\int x^{(2)}\phi(x)\nu(dx|x^{(1)} = y) = \int x^{(2)} \phi'(x)\nu(dx|x^{(1)} = y) \tag{2}$$

$$y \in H(\nu, \alpha), \quad \text{a.e.}(\nu) \quad .$$

(ii) Show that ϕ is admissible if and only if for some measurable functions C_i, γ_i, $i=1,2$,

$$\phi(x) = \begin{cases} 1 & \text{if} \quad x^{(1)} > C_2(x^{(2)}) \\ \gamma_2(x^{(2)}) & \text{if} \quad x^{(1)} = C_2(x^{(2)}) \\ 0 & \text{if} \quad C_1(x^{(2)}) < x^{(1)} < C_2(x^{(2)}) \\ \gamma_1(x^{(2)}) & \text{if} \quad x^{(1)} = C_1(x^{(2)}) \\ 1 & \text{if} \quad x^{(1)} < C_1(x^{(2)}) \end{cases} \quad . \tag{3}$$

[This is a continuation of 2.12.1 and 2.21.2.] (Matthes and Truax (1967).)

<u>7.12.2</u>

Prove that if ϕ is an admissible test and $Q \subset X$ with $\nu(Q) > 0$ then ϕ must also be admissible for the same problem with dominating measure $\nu_{|Q}$.

<u>7.12.3</u>

Let $X_1 = \bar{X}$ and $X_2 = S^2 + \bar{X}^2$ be the canonical statistics for the two-parameter exponential family generated by a $N(\mu, \sigma^2)$ random sample. (See Example 1.2.) Consider Figure 7.12.3. Draw the broken line parallel to $\mu_0 x_1 - x_2/2 = 0$ such that $\nu(R) = \nu(S)$. (ν is defined in Example 1.2.)

(i) Show that this is possible. (ii) Let ϕ_1 be the critical function for the test with acceptance region $Q' + R - S$, and let ϕ_0 be the critical function for the usual one-sided t-test, which has acceptance region $Q' = \{x_1 < 0$ or

$x_2 \geq cx_1^2\}$. Show

(1)
$$E_{(\mu,\sigma^2)}(\phi_1) < E_{(\mu,\sigma^2)}(\phi_0) \qquad \mu \leq 0$$
$$E_{(\mu,\sigma^2)}(\phi_1) > E_{(\mu,\sigma^2)}(\phi_0) \qquad \mu \geq \mu_0 \quad .$$

Hence ϕ_1 is a better test than ϕ_0 of

(2)
$$H_0\colon \ \mu \leq 0 \quad \text{versus} \quad H_1\colon \ \mu \geq \mu_0 \quad .$$

$[E(\phi_1 - \phi_0) = E(\chi_S - \chi_R)$. Now use Corollary 2.23.] (See Brown and Sackrowitz (1984). See also Exercise 7.14.6.)

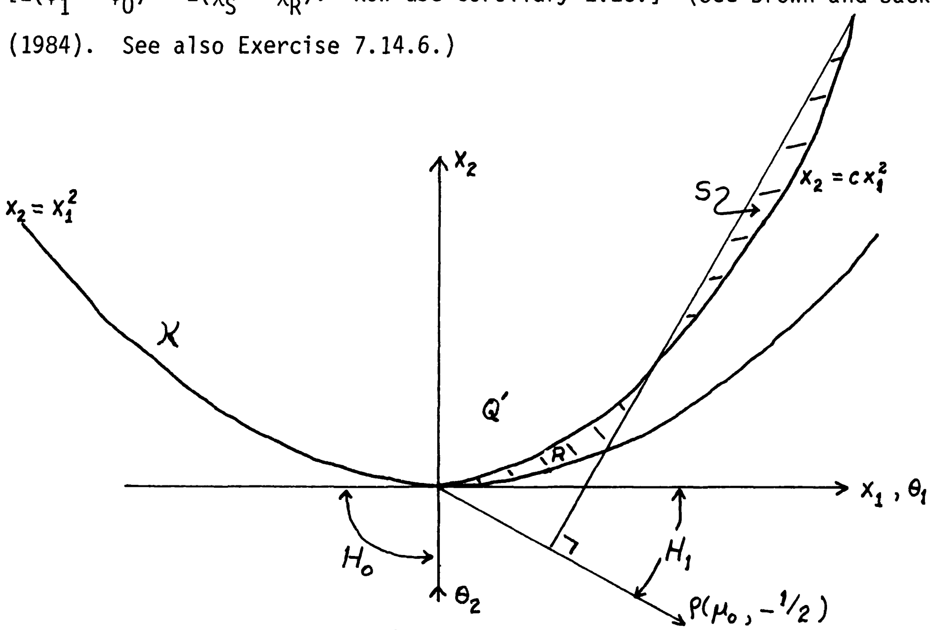


Figure 7.12.3: Diagram for Exercise 7.12.3

7.13.1

Here is an example which shows that something more than 7.13(1) is needed for validity of the conclusion of Lemma 7.13. Let $X \in R^2$ be bivariate $N(\theta, I)$. Consider the problem of testing $\Theta_0 = \{0\}$ versus $\Theta_1 = \{\theta\colon \theta_1 > 0, \theta_2 = -\theta_1^2\}$. Let $S = \{x \in R^2\colon \ x_2 \geq 0\}$.

(i) Show that $U = \phi$ but $U^* = (0, -1)$.

(ii) Verify that S satisfies 7.13(1) but not the remaining hypotheses of Lemma 7.13.

(iii) Let $\phi_1(x) = 1$ if $x \notin S$, $= 0$ otherwise. Show the conclusion of Lemma 7.13 does not apply to ϕ_1. [Let $\phi_2(x) = 1$ if $x_1 \geq 0$, $x_2 < \varepsilon$ or

$x_1 < 0$, $x_2 < -\varepsilon$. Show for $\varepsilon > 0$ sufficiently small ϕ_2 dominates ϕ_1.]

<u>7.13.2</u>

The additional assumptions of 7.13 are stronger than necessary. Let $X \sim N(\theta, I)$, $\Theta_0 = \{0\}$, S be as in 7.13.1. But now let $\Theta_1 = \{(\mu, \mu^4): \mu > 0\}$. Note that S satisfies 7.13(1) but does not satisfy either of the other two assumptions of Lemma 7.13. Show that if ϕ' is as good as ϕ then $\phi'(x) = 1$ for all $x \notin S$. Conclude that ϕ is admissible. [Show directly that if Q is an open set in S^c then

$$\lim_{\mu\to\infty} \frac{P_{(\mu,\mu^4)}(Q)}{P_{(\mu,\mu^4)}(S)} = \infty \quad .]$$

<u>7.14.1</u>

A test ϕ is said to have a nearly convex acceptance region if there is a closed convex set A such that $\phi(x) = 0$, $x \in A°$ and $\phi(x) = 1$ for $x \notin A$. (Thus, if ν is dominated by Lebesgue measure any test with nearly convex acceptance region is equivalent to one with a (closed) convex acceptance region. See the Remark following Corollary 4.17.) Suppose $\Theta_0 = \{\theta_0\}$ is simple in the setting of 7.12. Show that any Bayes test has nearly convex acceptance region.

<u>7.14.2</u>

Let ϕ_i be a sequence of critical functions with nearly convex acceptance regions. Suppose $\phi_i \to \phi$ weak* on L_∞. (See 4A.2(1) for the definition of weak* convergence.) Then ϕ has a nearly convex acceptance region. [Assume $\nu(R^k) < \infty$. To each ϕ_i there corresponds an A_i. Let $\{u_j\}$ be a countable dense subset of $\{u: ||u|| = 1\}$. Choose a subsequence $\{i'\}$ such that $\alpha_{A_{i'}}(u_j)$ converges for each u_j, say, $\alpha_{A_{i'}}(u_j) \to \alpha_j$. Let $A = \bigcap_j \bar{H}^-(u_j, \alpha_j)$. Then $\phi(x) = 0$, $x \in A°$ and $=1$ for $x \notin A$.]

<u>7.14.3</u>

Suppose $\Theta_0 = \{\theta_0\}$ is simple in the setting of 7.12.

(i) Show that the tests with nearly convex acceptance regions form a complete class.

(ii) Suppose, also, $\Theta_1 = R^k - \{\theta_0\}$ and ν is dominated by Lebesgue measure. Show that the tests with convex acceptance regions form a minimal complete class. [Use Theorem 4.14, 7.14.1, 7.14.2, and, for (ii), Theorem 7.14.]

<u>7.14.4</u>

Suppose the support of ν is a finite set, X. Let $\Theta_0 = \{\theta_0\} \in N = R^k$. (i) Prove that ϕ is admissible if and only if there is a closed convex set A such that $\phi(x) = 1$ if $X \notin A$, $= 0$ if $x \in A°$ or if $x \in$ r.i.F for some face F of A. (ii) Can you formulate an analogous complete class statement valid when X is countable and the assumptions of Theorem 6.23 are satisfied? [(i) Use Theorem 7.14, Corollary 7.10, and 7.12.2. (ii) Be careful; the characterization in (i) is not valid here, even when $X = \{0,1,\ldots\}^k$, and so will need to be modified.]

<u>7.14.5</u>

Consider a 2×2 contingency table. (See Exercise 1.8.1.) Two common tests for independence of row and column effects are the likelihood ratio test and the χ^2 test, based on the values of

$$\chi^2 = N \sum_{i,j} \frac{\left(Y_{ij} - \frac{Y_{i+} Y_{+j}}{N}\right)^2}{Y_{i+} Y_{+j}} .$$

(i) Use Theorem 7.14 to show that the χ^2 test is admissible.

(ii) Is the likelihood ratio test also admissible via Theorem 7.14?

(iii) Use 7.12.1 to prove both tests are admissible.

<u>7.14.6</u>

Show that the test with critical function ϕ_1 in Exercise 7.12.3 is admissible.

7.16.1

Let $X \in R^k$ be $N(\theta, I)$. Suppose $\Theta_0 = 0$ and $\Theta_1 = \{\theta: |\theta_i| > c \quad i=1,\ldots,k\}$. Consider level α tests of the form $\phi_1(x) = 1 - \chi_{\{t:|t_i|<a_1, i=1,\ldots,k\}}(x)$ and $\phi_2(x) = 1 - \chi_{\{||t||<a_2\}}(x)$. Note that ϕ_1 is admissible. Adjust k, c, α to provide an example where ϕ_2 dominates ϕ_1 except where π_{ϕ_1} is extremely small.

7.16.2

Consider the univariate linear model, as in 7.15. Show that the usual F test, 7.15(1), is Bayes. [Let $\eta \in R^s$. Let $\sigma^2 = 1/(1 + ||\eta||^2)$ and $\mu_i = \eta_i/(1 + ||\eta||^2)$, $i = r+1,\ldots,s$. Under Θ_1 also let $\mu_i = \eta_i/(1 + ||\eta||^2)$, $i=1,\ldots,r$. Under Θ_0 (resp. Θ_1) let η have density proportional to

$$(1 + ||\eta||^2)^{-p/2} \exp\left(\frac{||\eta||^2}{2(1 + ||\eta||^2)}\right)$$

$$\left(\text{resp.,} \quad (1 + ||\eta||^2)^{-p/2} \quad \exp\left(\sum_{r+1}^{s} \frac{\eta_i^2}{2(1 + ||\eta||^2)}\right)\right).]$$

(Kiefer and Schwartz (1965).)

7.16.3

Verify when r = 2 that the F test has the local optimality property described in 7.16(1). (This is called D-optimality.) [Write

$$\Sigma \frac{\partial}{\partial \mu_i^2} \pi_\phi(0, \sigma^2) = \Sigma \int \phi(y)\left(\frac{\partial}{\partial \mu_i^2} p_\mu(y)\Big|_{\mu=0}\right)dy$$

and use a general form of the Neyman-Pearson Lemma or Theorem 2.21.]

7.16.4

Let $X_1,\ldots,X_k$ be independent gamma variables with known indices $\alpha_1,\ldots,\alpha_k$ and unknown scale parameters $\sigma_1,\ldots,\sigma_k$. Consider the problem of testing the null hypothesis H_0: $\sigma_1 = \ldots = \sigma_k$. (In the special case where the X_i/σ_i are χ^2 variables resulting from a normal sample then this is the problem of testing *homogeneity of variance*. (In this notation the variances are $\sigma_1,\ldots,\sigma_k$.)) Show

(i) The likelihood ratio test for this problem has acceptance region

$$(1)\qquad S = \{x: \frac{(\Sigma\, x_i)^{\alpha_0}}{\Pi\, x_i^{\alpha_i}} \le C\}\ , \quad \text{where} \quad \alpha_0 = \sum_{i=1}^{k} \alpha_i \ .$$

(ii) When these distributions are written as a canonical exponential family the null hypothesis is linear in both parameter space and expectation space. Nevertheless, for $k \ge 3$, the acceptance region for the likelihood ratio test is not convex. (Hence there is no hope of proving its admissibility via Theorem 7.14.)

[(ii) Consider $k = 3$ and $\alpha_i \equiv \alpha$. Consider points of the form $x_z = (z, z, 1)$ on the boundary of the acceptance region S. Let $f(x) = \frac{\Pi x_i}{(\Sigma x_i)^3} - C$ so that $f(x) = 0$ for $x \in \partial S$. Show that for z sufficiently large $(\nabla f(x_z))'\ (D_2 f(x_z))(\nabla f(x_z)) < 0$.]

(iii) The likelihood ratio test is unique Bayes, hence admissible. Under H_1 let $\theta_i = 1/\sigma_i = (1 + \eta_i^2)$ where $\eta_i \in R$ are independent variables with density $|\eta_i|^{(\alpha_i-1)}\ (1 + \eta_i^2)^{-\alpha_i}$. Under H_0, $\theta_i = 1/\sigma_i \equiv (1 + \eta^2)$ where $\eta \in R$ has density $|\eta|^{(\alpha_0-1)}(1 + \eta^2)^{-\alpha_0}$. (This result is another one of many contained in Kiefer and Schwartz (1965).)

Note: It is not always true that a likelihood ratio test is admissible. For an interesting counter-example see Lehmann (1959, p.338) or Kiefer and Schwartz (1965, p.767).

7.17.1

Let $x \in R^2$ be bivariate normal, $N(\theta, I)$. Consider the problem of testing $\Theta_0 = \{0\}$ versus $\Theta_1 = \{\theta: \theta_1\theta_2 \ge 0\ ,\ \ ||\theta|| \ge 1\}$. Show that the non-randomized level $\alpha = .05$ test with acceptance region $\{x: ||x||^2 \le 5.991\}$ is inadmissible. (Can you also find a better test?) (Compare this result with 7.22.2 in which this test is admissible.)

7.17.2

Exercise 2.10.1 indicates a nontrivial testing problem where Θ_0 and Θ_1 are contiguous and all tests are admissible. Here is an example of the same phenomenon in which the null and alternative hypotheses are separated: Let $1 \le m < k$ and let $X = \{x \in R^k\colon x_i = 0 \text{ or } 1,\ i=1,\ldots,k,\ \Sigma x_i = m\}$. Let ν be counting measure on X, with $\{p_\theta\}$ the exponential family generated by ν. Let $\Theta_0 = \{0\}$, $\Theta_1 = \{\theta\colon ||\theta||^2 \ge 1\}$. (Other more restrictive definitions of Θ_1 will also suffice.) Let ϕ be any (possibly randomized) test. Then ϕ is admissible.

[It is possible to use Lemma 7.13 for this, but here is an easier argument. The aggregate family generated by $\{p_\theta\}$ contains $\{q_\xi\colon \xi \in X\}$ where $q_\xi(\cdot) = \chi_\xi(\cdot)$ and also $q_{\xi_0}(\cdot) \equiv \binom{k}{m}^{-1}$ where $\xi_0 = \xi(0)\ (\frac{m}{k})\underset{\sim}{1}$. If ϕ is inadmissible there exists a test ϕ' better than ϕ for testing Θ_0 versus Θ_1. Then (by continuity) ϕ' must be as good as ϕ for testing q_{ξ_0} versus $\{q_\xi\colon \xi \in X\}$. This implies $\phi'(x) \ge \phi(x)$, $x \in X$, and $\binom{k}{m}^{-1} \sum_{x \in X} \phi'(x) \le \binom{k}{m}^{-1} \sum_{x \in X} \phi(x)$.]

7.18.1

Let X_1, X_2 be independent gamma variables $\Gamma(\alpha_i, \lambda_i)$, $i=1,2$, variables with α_1, α_2 known. Consider the problem of testing $H_0\colon \lambda_1 = \lambda_2 = 1$ versus the alternative $H_1\colon \max_{i=1,2} |1 - \lambda_i| > \varepsilon$ for some given $\varepsilon > 0$. Show that any "intersection" test with acceptance region --

(1) $\phi(x) = 0$ iff $a_{i1} < x_i < a_{i2}$, $i=1, 2$, $(0 < a_{i1} < a_{i2} < \infty)$ --

is inadmissible. (See also 7.21.1.) [No admissible test can have an acceptance region with a sharp corner at $(x_1, x_2) = (a_{12}, a_{22})$ like (1) has. See Example 2.10.]

7.19.1

In Theorem 7.19 replace Θ_1^* by

(1) $\Theta_1^{**} = \{\theta_1 \in \bar{\Theta}_1 : \theta_1 \in N$ or there is a set $\{\theta_j' : j = 1, \ldots, J\} \subset N$ and a sequence $\{\zeta_i\} \subset \Theta_1$ with $\zeta_i \to \theta_1$ and $\{\zeta_i\} \subset \text{conhull}\,(\{\theta_j'\} \cup \{\theta_1\})\}$.

[Use 1.13.2.]

7.20.1

Prove the assertion in 7.20(3). [The extreme points of $\{J : J \in \mathcal{J}(\varepsilon), \int \theta J(d\theta) = v_0\}$, $v_0 \in \Theta$, are the distributions in this set which are concentrated on a single point; similarly the extreme points of $\{J : J \in \mathcal{J}(\varepsilon), \int \theta J(d\theta) = 0, \int \|\theta\|^2 J(d\theta) = \alpha\}$ are two-point distributions. The extreme points of $\bar{\Delta}(\varepsilon)$ are thus points (v, M) satisfying 7.19(2) with J either a one- or two-point distribution, as above. The extreme points of Δ are (contained in) the set of limits as $\varepsilon \to 0$ of these points.]

7.20.2

Prove the assertion following 7.20(4). [Let J be either a one- or two-point distribution.]

7.20.3

Generalize the assertion following 7.20(4) to apply to the situation where Θ is a twice differentiable manifold at θ_0. [First generalize 7.20(5)!]

7.21.1

In the setting of 7.18.1 consider the problem of testing $H_0 : \lambda_1 = \lambda_2 = 1$ versus the complementary alternative $H_1 : \lambda_1 \neq 1$ or $\lambda_2 \neq 1$. Show that the intersection test 7.18.1(1) is still inadmissible.

7.21.2

Consider the curved exponential family of Example 3.14 and 5.14. Let $\Theta_0 = \{\theta_0\}$ and $\Theta_1 = \Theta - \Theta_0$. To be specific take $\theta_0 = \theta(\lambda_0) = (-1,0)$; i.e., $\lambda_0 = 1$. One easily constructed test of Θ_0 is that which rejects when $|\hat{\lambda} - \lambda_0| > c_n$ with c_n chosen to give the desired level of significance. (Such a test can be constructed for any curved exponential family, and has certain asymptotic optimality properties as $n \to \infty$.) Show that for moderately large n and the usual levels of significance this test is inadmissible; although for every n there exists a (possibly very small) level of significance for which the test of this form is admissible. [Use 5.14 and Theorem 7.21. Except for small values of n or large values of c_n the acceptance region has a convex, but not strictly convex, form. Theorem 7.21 allows only very special admissible acceptance regions which are not strictly convex; and for appropriate values of n, c_n the above acceptance region is not of this special form.]

7.22.1

Let $X_1,\ldots,X_n$ be independent normal variables, $X_i \sim N(\mu, 1+\mu^2)$. Consider the problem of testing $H_0{:}\mu = 0$. Let $\phi_1 = 1$ if $|\bar{X}| > 1.96\ldots/\sqrt{n}$, $= 0$ otherwise; and $\pi_1(\mu) = E_\mu(\phi_1)$. Show

(i) ϕ_1 has level $\alpha = .05$ and is locally unbiased (i.e., $\pi_1'(0) = 0$, $\pi_1''(0) > 0$). (Is ϕ_1 also globally unbiased; i.e., $\pi_1(\mu) \geqq .05$??)

(ii) ϕ_1 is inadmissible. [Use 7.20(5) and Theorem 7.21. Note that $\theta_2 = -\frac{1}{2\sigma^2} = -(2(1+\mu^2))^{-1} \leqq -1/2$ to show ϕ_1 cannot satisfy 7.21(2) unless $H = 0$.]

(iii) Find a locally best locally unbiased level α test; i.e., the test which maximizes $\pi''(\mu)$ subject to $\pi(0) = \alpha$, $\pi'(0) = 0$. Use Theorem 7.22 to verify this test is admissible. [Admissibility actually follows directly from the fact that this test is the unique locally best locally unbiased level α test, but it may be instructive to note how this test can be written in the form 7.21(2) with $H = 0$.] Call this test ϕ_2.

((iv) Is ϕ_2 unbiased?? Is ϕ_2 better than ϕ_1?? If not, what is??)

(v) Generalize (i)-(iii) to arbitrary curved exponential families: Show that the locally unbiased test with parallel boundaries for the acceptance region is not locally best among locally unbiased tests unless $u_2 = 0$ in 7.20(5). State (convenient, frequently satisfied) conditions under which this parallel boundary test is inadmissible.

7.22.2

Let X be bivariate normal with mean θ and covariance I. Consider the problem of testing $\Theta_0 = 0$ versus $\Theta_1 = \{(\theta_1,\theta_2): \theta_1\theta_2 > 0\}$. Consider tests of the form $\phi(x) = \chi_{\{a(x_1+x_2)^2+b(x_1-x_2)^2 \geq c\}}(x)$, $a,b,c > 0$. (These tests are symmetric in (x_1,x_2).) Show that such a test is admissible if and only if $a \geq b$. The same result holds if $\Theta_1 = \{(\theta_1,\theta_2): \theta_1\theta_2 > 0,\ \theta_1^2+\theta_2^2 \leq 1\}$.