A NOTE ON THE NOTION OF ADEQUACY

by

D. Basu\textsuperscript{1} and S. C. Cheng\textsuperscript{2}

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

May, 1981
FSU Statistics Report M584

\textsuperscript{1}Research partially supported by NSF Grant No. 79-04693.

\textsuperscript{2}Department of Mathematics, Creighton University, Omaha, Nebraska 68178.
ABSTRACT

Written mainly for its mathematical interest this article studies in depth some aspects of the Skibinsky (1967) notion of adequacy in an abstract setting. Some known results on conditional independence are put together in Section II and then a basic factorization theorem is proved as Theorem H. This result throws some light on the factorization theorem for adequate subfields discussed in Section III. In Section IV the notion of minimum adequate subfield is studied in some details.
I. INTRODUCTION

Let $\theta \in \Theta$ be the parameter that defines the joint distribution $P_{\theta}$ of two random variates $X$ and $Y$, where $X$ is available for observation and $Y$ is a future random variate that needs to be predicted. In this context Skibinsky (1967) gave the following definition of an adequate or prediction sufficient statistic $T = T(X)$.

**DEFINITION.** The statistic $T = T(X)$ is adequate for $X$ with respect to $Y$ and $\theta$, in short, $T \text{ adq}(X; Y, \theta)$ if

(i) the conditional distribution of $X$, given $T$ and $\theta$, depends on $T$, $\theta$ only through $T$ [that is, $T = T(X)$ is a sufficient statistic when $X$ is regarded as the sample], and

(ii) the variates $X$ and $Y$ are conditionally independent, given $T$ and $\theta$, in short, $X \perp \! \! \! \perp Y|T$ with respect to $P_{\theta}$ for all $\theta \in \Theta$.

If we take a Bayesian view of the matter and regard $\theta$ as a random variate with a prior distribution $q$, then we restate the condition (i) as

(i') $X \perp \! \! \! \perp \theta|T$ irrespective of the particular choice of the prior distribution $q$.

It is easy to verify that the two conditions (i') and (i) are jointly equivalent to the single condition $X \perp \! \! \! \perp (Y, \theta)|T$. So we may restate the Skibinsky definition of adequacy in the following equivalent form:
DEFINITION. $T \operatorname{adq}(X; Y, \theta)$ if, irrespective of the specification of the prior $q$, $X \parallel (Y, \theta) | T$, that is, the posterior (predictive) distribution of $(Y, \theta)$, given $X$, depends on $X$ only through $T(X)$.

A wider and perhaps more natural definition of adequacy is

DEFINITION. The statistic $T = T(X)$ is adequate for $X$ with respect to $Y$ in the wide sense if, irrespective of the choice of the prior $q$, $X \parallel Y | T$, that is, the predictive distribution of $Y$, given $X$, depends on $X$ only through $T(X)$.

The two notions of adequacy have been studied in depth by Sugiura and Morimoto (1969), Ishii (1969), Lauritzen (1972), Takeuchi and Akahira (1975), Torgersen (1977), Cheng (1978), and others. The primary objective of this article is to study the Skibinsky notion of adequacy in an abstract mathematical setting and from the point of view of conditional independence. A basic factorization theorem (Theorem H of section II) for conditional independence is exploited to elucidate a factorization theorem for adequacy as stated by Sugiura and Morimoto (1969). The minimum adequate statistic is characterized in section IV.

In our abstract setting, the variables $X$ and $Y$ are idealized as subfields $\mathcal{B}$ and $\mathcal{C}$ in a basic measurable space $(X, \mathcal{A})$. The statistic $T$ is then a subfield $\mathcal{B}_0$ of $\mathcal{B}$. The expression "$X \parallel Y | T$ with respect to $P_\theta$" is rephrased as "$\mathcal{B} \parallel \mathcal{C} | \mathcal{B}_0$ with respect to $P_\theta$".
II. CONDITIONAL INDEPENDENCE

The basic statistical model is denoted by \((X, A, P)\), where \(X\) is the sample space, \(A\) a \(\sigma\)-field of subsets of \(X\), and \(P = \{P_\theta : \theta \in \Theta\}\) a family of probability measures on \(A\). Any sub-\(\sigma\)-field \(C\) of \(A\) will be referred to as subfield.

Given a family \(\{A_t : t \in T\}\) of measurable sets, we write \(\sigma(A_t : t \in T)\) for the subfield generated by the family of sets. Likewise, \(\sigma(f_t : t \in T)\) will stand for the smallest subfield \(C\) such that each \(f_t\) is \(C\)-measurable. By \(C \cup D\) we mean \(\sigma(C \cup D)\), that is the smallest subfield containing both \(C\) and \(D\).

A set \(N\) in \(A\) is \(P\)-null if \(P_\theta(N) = 0\) for all \(\theta \in \Theta\). Let \(\mathcal{N}\) denote the class of all \(P\)-null sets. For \(A, B\) in \(A\), we have \(A = B\) if \(A \Delta B\) is \(P\)-null. Similarly, for any two \(A\)-measurable functions, we write \(f = g\) if \(\{x : f(x) \neq g(x)\} \in \mathcal{N}\). Hereafter, we shall write \(f \in A\) for \(A\)-measurable function.

The completion \(\overline{C}\) of a subfield \(C\) is defined as \(\overline{C} = C \cup \sigma(N)\). Accordingly, a subfield \(C\) is called complete if \(N \subseteq C\). For any two subfields \(C\) and \(D\), by \(C \subseteq D\) we mean \(C \subseteq \overline{D}\). If \(C \subseteq D\) and \(D \subseteq C\), then \(C = D\). By a \(P\)-essentially \(C\)-measurable function, denoted by \(f \in \overline{C}\) or \(f \in C\), we mean that \(f\) is \(\overline{C}\)-measurable. A function \(f \in A\) is \(P\)-integrable if \(\int_X |f| \, dP_\theta < \infty\) for all \(\theta \in \Theta\). The restriction of \(P_\theta\) to a subfield \(C\) is denoted by \(P_\theta|_C\).
Let $(X, \mathcal{A})$ be a measurable space equipped with a fixed probability measure $P$. Given a function $f \in \mathcal{A}$, the function $f^* \in \mathcal{C}$ denotes a version of the conditional expectation $E(f|\mathcal{C})$ of $f$ given $\mathcal{C}$ with respect to $P$, that is, for all $C \in \mathcal{C}$,

$$\int f \, dP = \int f^* \, dP.$$

Two subfields $\mathcal{B}$ and $\mathcal{C}$ are said to be conditionally independent given the third subfield $\mathcal{D}$, denoted by "$\mathcal{B} \parallel \mathcal{C} | \mathcal{D}$", if

$$E(I_{B \cap C} | \mathcal{D}) = E(I_B | \mathcal{D})E(I_C | \mathcal{D}) \quad [P]$$

for all $B \in \mathcal{B}$ and $C \in \mathcal{C}$.

In this section we shall put together a number of results about the notion of conditional independence, which will be of central importance in our study.

**PROPOSITION A.** (Loéve, 1963) $\mathcal{B} \parallel \mathcal{C} | \mathcal{D}$ if and only if, for each $g \in \mathcal{C}$,

$$E(g | B \vee D) = E(g | \mathcal{D}) \quad [P]$$

**COROLLARY B.** $\mathcal{B} \parallel \mathcal{C} | \mathcal{D}$ if and only if, for each $g \in \mathcal{C}$, $E(g | B \vee D)$ has a $\mathcal{D}$-measurable version.

**PROPOSITION C.** The following statements are equivalent:

(i) $\mathcal{B} \parallel \mathcal{C} | \mathcal{D}$;

(ii) $\mathcal{B} \parallel \mathcal{C} \vee D | \mathcal{D}$;

(iii) $B \vee D \parallel \mathcal{C} \vee D | \mathcal{D}$.
PROPOSITION D. Let $\mathcal{D} \subseteq \mathcal{D}_1 \subseteq \mathcal{B}$. If $\mathcal{B} \parallel C | \mathcal{D}$, then $\mathcal{B} \parallel C | \mathcal{D}_1$.

Now suppose that $P_1 \ll P_0$ are two probability measures on a given measurable space $(X, \mathcal{A})$. Let $q = \frac{dP_1}{dP_0}$ denote a fixed version of the Radon-Nikodym derivative of $P_1$ with respect to $P_0$. Given a subfield $\mathcal{B}$ of $\mathcal{A}$, we write $\frac{dP_1}{dP_0} | \mathcal{B}$ to denote the Radon-Nikodym derivative of $P_1 | \mathcal{B}$ with respect to $P_0 | \mathcal{B}$. By noting that, for all $B \in \mathcal{B}$,

$$\int_B E_0(q|\mathcal{B})dP_0 = \int_B qdP_0 = P_1(B)$$

we obtain the following lemma immediately.

**LEMMA E.** $E_0(q|\mathcal{B})$ is a version of $\frac{dP_1}{dP_0} | \mathcal{B}$, i.e., they are $P_0$-equivalent.

**COROLLARY F.** Let $q_1$ be a fixed version of $E_0(q|\mathcal{B})$ and let $\mathcal{D}$ be a subfield of $\mathcal{B}$. Then, for all $f \in \mathcal{B}$,

$$E_0(fq_1|\mathcal{D}) = E_0(fq|\mathcal{D}) \quad [P_0]$$

**PROOF.** Let $f \in \mathcal{B}$ be given. It follows from Lemma E that, for all $D \in \mathcal{D}$,

$$\int_Dfq_1dP_0 = \int_DfdP_1 = \int_DfqdP_0.$$
Given a subfield \( B \) of \( A \), let
\[
B_0 = \{x : E_0(q|B)(x) > 0\}
\]
it is clear that \( P_1(B_0') = 0 \). An important relation between the two conditional expectations (given the subfield \( B \)) under \( P_0 \) and \( P_1 \) is as follows:

**Lemma C.** Let \( f \in A \). Then
\[
E_1(f|B) = \frac{E_0(f|B)}{E_0(q|B)} \quad [P_1]
\]

**Proof.** Let \( f \in A \) be given. The proof consists of verifying the following equalities:

\[
\int_B \frac{E_0(fq|B)}{E_0(q|B)} \ dP_1
\]

\[
= \int_{B|B_0} \frac{E_0(fq|B)}{E_0(q|B)} \ dP_1 \quad [\text{where } B_0 \text{ is defined as in (a)}]
\]

\[
= \int_{B|B_0} \frac{E_0(fq|B)}{E_0(q|B)} \ E_0(q|B) \ dP_0 \quad [\text{by Lemma E}]
\]

\[
= \int_{B|B_0} E_0(fq|B) \ dP_0
\]

\[
= \int_{B|B_0} fq \ dP_0
\]

\[
= \int_{B|B_0} f \ dP_1 \quad [\text{since } P_1 << P_0]
\]

\[
= \int_B f \ dP_1 \quad [\text{since } P_1(B_0') = 0].
\]
Now choose and fix two subfields $B$ and $C$ of $A$, and let $D$ be a non-trivial common subfield of both $B$ and $C$. As in Corollary $F$, $q_1$ denotes a fixed version of $E_0(q|B)$. Likewise, $q_2$ and $q_3$ represent fixed versions of $E_0(q|C)$ and $E_0(q|D)$ respectively. Define a $C$-measurable function $q_4$ as follows:

$$(b) \quad q_4 = \begin{cases} 
q_2 / q_3 & \text{if } q_3 > 0 \\
0 & \text{if } q_3 = 0
\end{cases}$$

Let $D = \{ x : q_3(x) > 0 \}$. It is easy to see that $P_1(D') = 0$. We are now in a position to prove the following factorization theorem for conditional independence.

**Theorem H.** Suppose $A = B \lor C$, $P_1 \ll P_0$, and $B \parallel C|D$ with respect to $P_0$. Then a necessary and sufficient condition in order that $B \parallel C|D$ with respect to $P_1$ is that

$$(c) \quad q = q_1 q_4 \quad [P_0]$$

where $q_1$ and $q_4$ are defined as above.

**Proof.** The "necessary" part is proved as follows. Suppose that $B \parallel C|D$ with respect to $P_1$. Let $f \in B$ and $g \in C$ be given, we have

$$\int_X fgq_1 q_4 dP_0 = \int_D fgq_1 q_2^{-1} q_3 dP_0 \quad \text{[by (b) and the fact that]}$$

$$D = \{ x : q_3(x) > 0 \}$$
\[ \int_{D} E_{0}(fgq_{1}q_{2} \mid D)q_{3}^{-1}dP_{0} \quad \text{[since } q_{3} \in D] \]

\[ \int_{D} E_{0}(fq_{1} \mid D)E_{0}(gq_{2} \mid D)q_{3}^{-1}dP_{0} \quad \text{[since } B \parallel C \mid D \text{ with respect to } P_{0}] \]

\[ \int_{D} E_{0}(fq \mid D)E_{0}(gq \mid D)q_{3}^{-1}dP_{0} \quad \text{[by Corollary F]} \]

\[ \int_{D} E_{0}(fq \mid D) \frac{E_{0}(gq \mid D)}{q_{3}} dP_{1} \quad \text{[by Lemma E]} \]

\[ \int_{D} E_{1}(f \mid D)E_{1}(g \mid D) dP_{1} \quad \text{[by Lemma G]} \]

\[ \int_{D} E_{1}(fg \mid D) dP_{1} \quad \text{[since } B \parallel C \mid D \text{ with respect to } P_{1}] \]

\[ \int_{D} fg dP_{1} \]

\[ \int_{X} fg dP_{1} \quad \text{[since } P_{1}(D) = 1] \]

\[ \int_{X} fgqdP_{0}. \]

Since \( A = B \lor C = \sigma\{fg : f \in B \text{ and } g \in C\} \), equation (c) follows at once as required.

The "sufficient" part is proved as follows. Suppose that equation (c) holds. Let \( f \in B \) and \( g \in C \) be given. Then,
\[
E_1(fg|D) = \frac{E_0(fgq_4|D)}{q_3} \quad [P_1] \quad \text{[by Lemma G]}
\]
\[
= q_3^{-2}E_0(fgq_1q_2|D) \quad [P_1] \quad \text{[by (b) and the fact that \(P_1(D') = 0\)]}
\]
\[
= E_0(fg_1|D)E_0(gq_2|D) \quad [P_0] \quad \text{[since \(B|C|D\) with respect to \(P_0\)]}
\]
\[
= E_1(\ell|D)E_1(B|D) \quad [P_1] \quad \text{[by Lemma G and Corollary F]}
\]

Hence, the proof is complete.

**REMARK.** If the assumption \(D \subset B \cap C\) is omitted, by replacing \(B\) by \(B \setminus D\) and \(C\) by \(C \setminus D\) in the above proof, Theorem H is easily modified to hold in the light of (iii) of Proposition C.

Now, by letting \(D\) be the trivial subfield \(\{\phi, X\}\), we obtain the following immediate consequence.

**COROLLARY I.** Suppose \(A = B \setminus C, P_1 \ll P_0\), and \(B\) and \(C\) are independent with respect to \(P_0\). Then, \(B\) and \(C\) are independent with respect to \(P_1\) if and only if
\[
q = q_1q_2 \quad [P_0]
\]
where \(q_1\) and \(q_2\) are defined as before.
III FACTORIZATION THEOREM

Let \((X, A, P)\), where \(P = \{P_\theta : \theta \in \Theta\}\), be a statistical model. Throughout this paper, \(B\) and \(C\) will denote two subfields of \(A\), and \(B_0\) will denote a subfield of \(B\). The subfield \(B_0\) is said to be sufficient for \(B\) with respect to \(P\) if, for each \(f \in B\), there exists a \(P\)-integrable function \(f^* \in B_0\) such that
\[
f^* = E_\theta(f|B_0)P_\theta\]
for all \(\theta \in \Theta\). We shall write \("B_0 \text{ suff}(B, P)"\) to indicate that \(B_0\) is sufficient in the above sense.

**DEFINITION** (Skibinsky, 1967). \(B_0\) is said to be adequate (or prediction sufficient) for \(B\) with respect to \(C\) and \(P\) if

(i) \(B_0 \text{ suff}(B, P)\);

(ii) \(\frac{B_0 \parallel C}{B_0}\) for all \(\theta \in \Theta\) i.e., \(B\) and \(C\) are conditionally independent given \(B_0\) with respect to each \(P_\theta\).

We shall also write \("B_0 \text{ adq}(B; C, P)"\) to mean that \(B_0\) is adequate for \(B\) in the sense of Skibinsky.

For each pair \(C \in C\) and \(\theta \in \Theta\) with \(P_\theta(C) > 0\), define a new probability measure \(P_\theta^C\) on \(A\) as follows:

\[
P_\theta^C(A) = \frac{P_\theta(AC)}{P_\theta(C)} \quad \text{for all } A \in A.
\]

Let \(P(C)\) denote the family of probability measures on \(A\) thus defined. Clearly, \(P_\theta^X = P_\theta\) for all \(\theta \in \Theta\).

Thus, \(P \subseteq P(C)\).
THEOREM A. (Skibinsky, 1967). The following statements are equivalent:

(i) $B_0$ suff$(B, P(C))$.
(ii) $B_0$ adq$(B; C, P)$.
(iii) Corresponding to each $f \in B$, there exists a $P$-integrable function $f^+ \in B_0$ such that

$$f^+ = E_\theta(f|B_0 \cap C) \quad [P_\theta] \quad \text{for all } \theta \in \Theta.$$

It is assumed hereafter that the family $P$ is dominated by some $\sigma$-finite measure $\lambda$ on $(X, A)$, and that $A$ is the smallest $\sigma$-field containing both $B$ and $C$, that is, $A = B \cap C$.

Let us first formulate a factorization criterion for the self-dominated case — $P \ll P_{\theta_0}$ where $\theta_0 \in \Theta$. We denote by $q_\theta = dP_\theta/dP_{\theta_0}$ a fixed version of the Radon-Nikodym derivative of $P_\theta$ with respect to $P_{\theta_0}$.

THEOREM B. Suppose $P$ is self-dominated by $P_{\theta_0}$ for $\theta_0 \in \Theta$. A necessary and sufficient condition for $B_0$ adq$(B; C, P)$ is that, for each $\theta \in \Theta$, $q_\theta$ is $P$-essentially $B_0 \cap C$-measurable.

PROOF. It follows from Theorem A that $B_0$ adq$(B; C, P)$ iff, for each $f \in B$, there corresponds a $P$-integrable function $f^+ \in B_0$ such that

$$f^+ = E_\theta(f|B_0 \cap C) \quad [P_\theta] \quad \text{for all } \theta \in \Theta,$$
that is,

\[ \int f dP_\Theta = \int f^+ dP_\Theta \]

for all \( D \in \mathcal{B}_0 \vee C \) and \( \Theta \in \Theta \). Since \( P \ll P_{\Theta_0} \), we at once have

\[ \int f q_\Theta dP_{\Theta_0} = \int f^+ q_{\Theta} dP_{\Theta_0} \]

for all \( D \in \mathcal{B}_0 \vee C \) and \( \Theta \in \Theta \). By noting that \( f^+ \) is a version of \( E_{\Theta_0}(f|\mathcal{B}_0 \vee C) \), and applying the self-adjoint property of conditional expectations, the left-hand side of the above equality becomes

\[ \int f^+ q_{\Theta} dP_{\Theta_0} = \int E_{\Theta_0}(f|\mathcal{B}_0 \vee C) q_{\Theta} dP_{\Theta_0} \]

\[ = \int f E_{\Theta_0}(q_{\Theta} | \mathcal{B}_0 \vee C) dP_{\Theta_0} \]

Thus, \( \mathcal{B}_0 \) adq(\( \mathcal{B}; C, P \)) iff

\[ \int f q_{\Theta} dP_{\Theta_0} = \int f E_{\Theta_0}(q_{\Theta} | \mathcal{B}_0 \vee C) dP_{\Theta_0} \]

for all \( f \in \mathcal{B}, D \in \mathcal{B}_0 \vee C, \Theta \in \Theta \). Now taking

\[ D = C \in C \text{ and } f = I_B \in \mathcal{B}. \]

we get

12
\[ \int_{BC} q_\theta dP_\theta = \int_{BC} E_\theta (q_\theta | B_0 \forall C) dP_\theta \]

for all \( B \in B, C \in C, \theta \in \Theta \). Since \( A = B \forall C = \sigma(BC: B \in B, C \in C) \), it follows that

\[ \int_{A} q_\theta dP_\theta = \int_{A} E_\theta (q_\theta | B_0 \forall C) dP_\theta \]

for all \( A \in A \) and \( \theta \in \Theta \). This implies that

\[ q_\theta = E_\theta (q_\theta | B_0 \forall C) \quad [P_\theta] \quad \text{for all } \theta \in \Theta. \]

Consequently, \( B_0 \) adq(\( B; C, P \))

iff \( q_\theta = E_\theta (q_\theta | B_0 \forall C) \quad [P_\theta] \quad \text{for all } \theta \in \Theta \)

iff \( q_\theta = E_\theta (q_\theta | B_0 \forall C) \quad [P] \quad \text{for all } \theta \in \Theta \)

iff \( q_\theta \in B_0 \forall C \quad [P] \quad \text{for all } \theta \in \Theta. \)

Let \( A_\alpha \) be the subfield generated by the functions \( q_\theta, \theta \in \Theta \), namely, \( A_\alpha = \sigma(q_\theta : \theta \in \Theta) \). The above theorem may be stated in following alternative form.

**COROLLARY C.** Suppose \( P \) is self-dominated by \( P_\theta \) for some \( \theta_0 \in \Theta \). Then \( B_0 \) adq(\( B; C, P \)) if and only if \( A_\alpha \subset B_0 \forall C \quad [P]. \)
**Lemma D.** \( E_0 \text{ adq}(\mathcal{B}; C, P) \) if and only if \( E_0 \text{ adq}(\mathcal{B}; \mathcal{B}, \mathcal{P}) \), where \( \mathcal{P} \) is the convex hull of \( P \).

**Proof** Noting that \( P \subset \mathcal{P} \), the "if" part follows immediately. To prove the "only if" part, choose and fix \( Q \in \mathcal{P} \). Then,

\[
Q = \Sigma a_1 P_{\theta_1}
\]

where \( \theta_1 \in \Theta \), \( a_1 > 0 \), and \( \Sigma a_1 = 1 \).

Since \( E_0 \text{ adq}(\mathcal{B}; C, P) \), it follows from (iii) of Theorem A that, for each \( f \in \mathcal{B} \), there exists \( f^+ \in E_0 \) such that \( f^+ = E_0(f|B_0 \cup C) \) \([P_\theta]\) for all \( \theta \in \Theta \). Thus, for all \( D \in E_0 \cup C \) and \( \theta \in \Theta \),

\[
\int f dP_\theta = \int f^+ dP_\theta
\]

and, hence,

\[
\int f dQ = \Sigma a_1 \int f dP_{\theta_1} = \Sigma a_1 \int f^+ dP_{\theta_1} = \int f^+ dQ.
\]

That is, \( f^+ = E_0(f|B_0 \cup C) \) \([Q]\). Therefore, \( E_0 \text{ adq}(\mathcal{B}; C, \mathcal{P}) \).

Now, returning to the case where \( P \ll \lambda \), it is known that there exists a countable collection \( \{P_{\theta_1} : i = 1, 2, \ldots\} \) of measures in \( P \) such that the convex combination

(a) \( Q = \Sigma a_1 P_{\theta_1} \), where \( a_1 > 0 \) and \( \Sigma a_1 = 1 \),

dominates the family \( P \). Consider the new family \( P' = P \cup \{Q\} \).

Clearly, \( P' \) is self-dominated by \( Q \). On account of Lemma D and the fact that \( P \subset P' \), Theorem B and Corollary C have their
counterparts in the dominated case. That is,

**COROLLARY E.** Let $P \ll \lambda$. A necessary and sufficient condition for $\mathcal{B}_0 \adq(\mathcal{B}; C, P)$ is that, for each $\theta \in \Theta$, $dP_\theta/dQ$ is $P$-essentially $\mathcal{B}_0 \nu C$-measurable, where $Q$ is defined as in (a).

**COROLLARY F.** Let $P \ll \lambda$. Then, $\mathcal{B}_0 \adq(\mathcal{B}; C, P)$ if and only if $\sigma(dP_\theta/dQ : \theta \in \Theta) \subseteq \mathcal{B}_0 \nu C$ [P], where $Q$ is defined as in (a).

Now we are ready to prove our main result — the factorization theorem for adequate subfields. In the following theorem, denote by $p_\theta = dP_\theta/d\lambda$ a fixed version of the Radon-Nikodym derivative of $P_\theta$ with respect to the dominating $\sigma$-finite measure $\lambda$.

**THEOREM G.** Suppose $P \ll \lambda$. If there exist a nonnegative function $h \in \mathcal{B}$ and a collection $\{g_\theta : \theta \in \Theta\}$ of nonnegative $\mathcal{B}_0 \nu C$-measurable functions such that

$$p_\theta = g_\theta h \quad [\lambda] \quad \text{for all } \theta \in \Theta,$$

then $\mathcal{B}_0 \adq(\mathcal{B}; C, P)$.

If, in addition, $\mathcal{B} || C | \mathcal{B}_0$ with respect to $\lambda$ and $\lambda$ is a unitary measure, then the reverse implication also holds.

**PROOF.** Suppose that the factorization holds. Let $Q = \Sigma a_i P_\theta^i$ be defined as in (a). Then

$$\frac{dQ}{d\lambda} = h \Sigma a_i g_\theta^i \quad [\lambda]$$

15
For each $\theta \in \Theta$, define $q_\theta = \frac{g_\theta}{\sum a_1 g_1} E$ where

$$E = \{ x : \sum a_1 g_1(x) > 0 \}. \text{ Clearly, } Q(E') = 0 \text{ and } q_\theta \in B_0 \nu C \text{ for all } \theta \in \Theta. \text{ In accordance with Corollary E, the assertion }$$

"$B_0 \text{ adq}(B; C, P)$" follows if we can prove that $q_\theta$ is a version of $dP_\theta/dQ$ for each $\theta \in \Theta$. To see this, observe that, for all $A \in \Lambda$,

$$\int_A q_\theta dQ = \int_A \frac{g_\theta}{\sum a_1 g_1} dQ = \int_A \frac{g_\theta}{\sum a_1 g_1} \frac{h \sum a_1 g_1}{h} d\lambda$$

$$= \int_A g_\theta h d\lambda = P_\Theta(AE) = P_\Theta(A),$$

the last equality following from the fact that $E'$ is $Q$-null and therefore is $P_\Theta$-null.

For the converse, if $B_0 \text{ adq}(B; C, P)$, then it follows from Corollary E that every $dP_\theta/dQ$ is $P$-essentially $B_0 \nu C$-measurable, where $Q$ is defined as in (a). Since $P \ll \lambda$, it is clear that

$$P_\Theta = \frac{dP_\Theta}{dQ} \frac{dQ}{d\lambda} \quad [\lambda]$$

In view of Theorem H of section 2, $dQ/d\lambda$ can be factored as

$$\frac{dQ}{d\lambda} = hh_1 \quad [\lambda],$$

where $h \in \Theta$ and $h_1 \in B_0 \nu C$. Letting $g_\theta = h_1 \frac{dP_\theta}{dQ}$, we at once have the factorization as required.
Now consider a product statistical model \((X_1 \times X_2, B \times C, P)\).
Let \(B_0\) denote a subfield of \(B\). Let \(B, B_0\), and \(C\) also denote the corresponding (marginal) subfields of \(B \times C\) in the product space.

Suppose that \(P \ll \lambda = \lambda_1 \times \lambda_2\), where \(\lambda_1\) and \(\lambda_2\) are \(\sigma\)-finite measures on \((X_1, B)\) and \((X_2, C)\), respectively. It is known that there exists a unitary measure \(\bar{\lambda}_i (i = 1, 2)\) such that \(\lambda_i = \bar{\lambda}_i (i = 1, 2)\). Let \(\bar{\lambda} = \bar{\lambda}_1 \times \bar{\lambda}_2\). Then \(\lambda \equiv \bar{\lambda}\) and so \(P \ll \bar{\lambda}\).

Since \(B \parallel C\) with respect to \(\bar{\lambda}\), it follows that \(B \parallel C | B_0\) with respect to \(\bar{\lambda}\). As an immediate consequence of Theorem \(G\), we have the following result.

**Corollary H.** Under the foregoing assumptions, a necessary and sufficient condition for \(B_0 \text{ adq}(B; C, P)\) is that there exist a nonnegative function \(h \in B\) and a collection \(\{g_\theta : \theta \in \Theta\}\) of nonnegative \(B_0 \vee C\)-measurable functions such that, for all \(\theta \in \Theta\),
\[
dP_\theta / d\lambda = g_\theta h \quad [\bar{\lambda}].
\]

Let us now illustrate the factorization theorem by the following examples.

**Example.** Let \(X(1) < X(2) < \ldots < X(n)\) denote the order statistics of a random sample of size \(n\) from a population with density from an exponential family
\[
p(x|\theta) = B(\theta)h(x) Q(\theta)R(x), \quad \theta \in \Theta.
\]
Consider the problem of predicting \(X(m)\) after observing
The point density of $X(1), X(2), \ldots, X(r)$, where $1 \leq r \leq m \leq n$. The point density of $X(1), X(2), \ldots, X(r)$, and $X(m)$ is given by

$$P_B(X(1), X(2), \ldots, X(r), X(m)) = \frac{n! \left[B(\theta)\right]^n}{(m-r-1)!(n-m)!}.$$

$$Q(\theta)\left[\sum_{i=1}^{r} R(x(i)) + R(x(m))\right] = \frac{X(m)}{X(r)} \left[\int h(t) e^{Q(\theta)R(t)} dt\right]^{n-r-1}.$$

$$\left[\int_{X(m)}^{\infty} h(t) e^{Q(\theta)R(t)} dt\right]^{n-m} = \left[\prod_{i=1}^{r} h(x(i)) \cdot h(x(m))\right].$$

Let

$$G_B(\sum_{i=1}^{r} R(x(i)), x(r), x(m)) = \frac{n! \left[B(\theta)\right]^n}{(m-r-1)!(n-m)!}.$$

$$Q(\theta)\left[\sum_{i=1}^{r} R(x(i)) + R(x(m))\right] = \frac{X(m)}{X(r)} \left[\int h(t) e^{Q(\theta)R(t)} dt\right]^{n-r-1}.$$

$$\left[\int_{X(m)}^{\infty} h(t) e^{Q(\theta)R(t)} dt\right]^{n-m} = h(x(m))$$

and

$$H(x(1), \ldots, x(r)) = \prod_{i=1}^{r} h(x(i)).$$

Then it follows from the factorization theorem that
\( T \sim \text{adq}(X; X(m), P) \), where \( T = \left( \sum_{i=1}^{r} R(X(i)), X(r) \right) \)

and \( X = (X(1), \ldots, X(r)) \).

**EXAMPLE.** Let \( X_1, X_2, \ldots, X_n \) be \( n \) independent observable random variables from \( N(\theta, 1) \), \(-\infty < \theta < \infty\), and let \( Y = \bar{X} + Z \), where \( \bar{X} \) is the sample mean of \( X = (X_1, X_2, \ldots, X_n) \) and \( Z \) is a not yet observed random variable that is independent of \( \bar{X} \). The joint density of \( \bar{X} \) and \( Y \) is readily found to be

\[
 p_{\bar{X}, Y}(x, y) = \left( \frac{1}{\sqrt{2\pi}} \right)^{n+1} e^{-\frac{1}{2} \sum_{i=1}^{n} (x_i - \bar{x})^2 - \frac{1}{2}(y-x)^2} \\
= \left( \frac{1}{\sqrt{2\pi}} \right)^{n+1} e^{-\frac{1}{2} \sum_{i=1}^{n} x_i^2 + n\bar{x} - \frac{1}{2} n\theta^2 - \frac{1}{2}(y-x)^2}
\]

Let \( g_{\theta}(\bar{x}, y) = e^{n\bar{x}} - \frac{1}{2} n\theta^2 - \frac{1}{2}(y-x)^2 \)

and

\[
h(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^n e^{-\frac{1}{2} \sum_{i=1}^{n} x_i^2}
\]

In view of Theorem G it is known that \( \bar{X} \sim \text{adq}(X; Y, P) \).

**EXAMPLE.** Let \( X_1, X_2, \ldots, X_n \) and \( Y \) be \( n+1 \) random variables whose joint density is given as follows.
\[ p_\theta(x_1, \ldots, x_n, y) = \frac{\theta^n (x_1 \cdots x_n)^{\theta-1}}{\pi(1+y^2)} \{1 + \frac{\theta}{\pi} [2(x_1 x_2 \cdots x_n)^{\theta}-1] \tan^{-1} y\} \]

where \(0 < \theta < 1\), \(-\infty < y < \infty\), and \(0 \leq x_i \leq 1\) for \(i = 1, 2, \ldots, n\).

Let \(X = (X_1, X_2, \ldots, X_n)\) and \(T = X_1 X_2 \cdots X_n\). Define

\[ g_\theta(t, y) = \frac{\theta^n t^{\theta-1}}{1+y^2} \left[1 + \frac{\theta}{\pi} (2t^{\theta}-1) \tan^{-1} y\right] \]

and

\[ h(x) = \frac{1}{\pi}. \]

Then, it follows from Theorem G that \(T \text{ adq}(X; Y, P)\).
IV MINIMUM ADEQUACY.

It is well known that a minimum sufficient subfield always exists in the dominated statistical models. Thus, under the assumption of dominance, Skibinsky's theorem (Theorem A of Section III) assures the existence of a minimum adequate subfield. More explicitly, in this section, we shall give a construction of a minimum adequate subfield by the factorization criterion.

Consider the class of all adequate subfields (for \( \mathcal{B} \) with respect to \( C \) and \( P \)) equipped with the partial ordering "\( \subset [P] \)". Its minimal and minimum elements are given as follows.

DEFINITION. A subfield \( \mathcal{B}^* \) of \( \mathcal{B} \) is said to be minimal adequate for \( \mathcal{B} \) with respect to \( C \) and \( P \) if

(i) \( \mathcal{B}^* \) adq\( \mathcal{B}; C, P \);

(ii) \( \mathcal{B}_0 \) adq\( \mathcal{B}; C, P \) and \( \mathcal{B}_0 \subset \mathcal{B}^* [P] \) imply that \( \mathcal{B}^* \subset \mathcal{B}_0 [P] \).

DEFINITION. A subfield \( \mathcal{B}^* \) of \( \mathcal{B} \) is said to be minimum adequate for \( \mathcal{B} \) with respect to \( C \) and \( P \) if

(i) \( \mathcal{B}^* \) adq\( \mathcal{B}; C, P \);

(ii) \( \mathcal{B}_0 \) adq\( \mathcal{B}; C, P \) implies that \( \mathcal{B}^* \subset \mathcal{B}_0 [P] \).

It should be noted that if a statistical model has a minimum adequate subfield, then that subfield is also minimal adequate. Whether or not we should distinguish "minimal adequacy" from "minimum adequacy" hardly matters in the light of the following
result [which is a counterpart of Corollary 3 of Burkholder (1961)].

**PROPOSITION A.** If $B^*$ is minimal adequate, then it is also minimum adequate.

**PROOF.** Let $B_0$ be an arbitrary adequate subfield of $B$. Applying Theorem 4 of Burkholder (1961) in conjunction with Skibinsky's theorem, we know that $\overline{B}^* \cap \overline{B}_0 \text{adq}(B; C, P)$. Owing to the minimality of $B^*$, it follows that

$$B^* \subset (\overline{B}^* \cap \overline{B}_0) \lor \sigma(N) = \overline{B}_0 \lor \sigma(N) = \overline{B}_0.$$  

This proves that $B^*$ is minimum adequate.

We now write "$B^*$ min adq($B; C, P$)" to indicate that $B^*$ is minimum (or minimal) adequate in the above sense. In general, there may not exist a minimum adequate subfield. Pitcher's example (1957) can be easily extended to be such a counterexample.

Suppose $P << \lambda$ where $\lambda$ is some $\sigma$-finite measure on $A$. Let $Q$ be a convex combination of measures from $P$ such that $P << Q$, and let $q_0$ be a fixed version of $dP_0/dQ$ for all $\theta \in \Theta$. For each $\theta \in \Theta$ and each $P$-integrable function $g \in C$, define a function $g_0 \in B$ by
where $D = \{x : E_Q(q_\theta | B)(x) > 0\}$. Clearly, $P_\theta(D') = 0$.

Let $B_\star$ be the subfield of $B$ that is induced by such functions $g_\theta$, where $\theta \in \Theta$ and $g \in C$. Now note that, for each $g \in C$, $E_\theta(g|B)$ has $B_\star$-measurable version $g_\theta$. Thus we at once have the following result.

**Lemma B.** $B \| C | B_\star$ with respect to $P_\theta$ for all $\theta \in \Theta$.

Moreover, $B_\star$ is a subfield of every adequate subfield.

That is,

**Lemma C.** If $B_\star$ adq$(B; C, P)$, then $B_\star \subset E_0$ [P].

**Proof.** It suffices to show that, for each $\theta \in \Theta$ and each $g \in C$, there exists a function $h_\theta \in E_0$ such that $g_\theta = h_\theta$ [P].

Choose and fix $\theta \in \Theta$ and $g \in C$. Define a function $h_\theta \in E_0$ as follows:

\[(b) \quad h_\theta = \begin{cases} \quad E_0(g|B_0) & \text{on } D_0 \\ 0 & \text{on } D'_0 \end{cases} \]

where $D_0 = \{x : E_Q(q_\theta | B_0)(x) > 0\}$. It is clear that $P_\theta(D'_0) = 0$.

Since $E_Q(q_\theta | B)$ coincides with the Radon-Nikodym derivative of $P_\theta|B$ with respect to $Q|B$, it follows that, for all $B \in B$,
\[
\int_{B} g_{\theta} F_{Q}(q_{\theta} | \mathcal{S}) dQ = \int_{B} g_{\theta} dP_{\theta}
\]

\[
= \int_{B} E_{\theta}(g | \mathcal{S}) dP_{\theta} \quad \text{[by (a)]}
\]

\[
= \int_{B} E_{\theta}(g | \mathcal{S}) dP_{\theta} \quad \text{[since } B \parallel C | \mathcal{D} \text{ with respect to } P_{\theta}\text{]}
\]

\[
= \int_{B} h_{\theta} dP_{\theta} \quad \text{[by (b)]}
\]

\[
= \int_{B} h_{\theta} F_{Q}(q_{\theta} | \mathcal{S}) dQ
\]

This yields that

\[
g_{\theta} F_{Q}(q_{\theta} | \mathcal{S}) = h_{\theta} F_{Q}(q_{\theta} | \mathcal{S}) \quad [Q]
\]

But \( P \equiv Q \) and \( F_{Q}(q_{\theta} | \mathcal{S}) > 0 \) on \( D \). Hence, we get \( g_{\theta} = h_{\theta} [P] \) on \( D \).

It remains to show that \( g_{\theta} = h_{\theta} [P] \) on \( D' \). Recall that
\( g_{\theta} = 0 \) on \( D' = \{ x : E_{Q}(q_{\theta} | \mathcal{S})(x) = 0 \} \) and \( h_{\theta} = 0 \) on \( D' = \{ x : E_{Q}(q_{\theta} | \mathcal{S})(x) = 0 \} \). The desired result follows if we can prove that

\[
(c) \quad E_{Q}(q_{\theta} | \mathcal{S}) = E_{Q}(q_{\theta} | \mathcal{S}) \quad [Q]
\]
Since $\mathcal{B}_0 \suff(\mathcal{B}, P)$, it is known that $\mathcal{B}_0 \suff(\mathcal{B}, P \cup \{Q\})$.
Choose and fix $f \in \mathcal{B}$. There exists a $f^+ \in \mathcal{B}_0$ such that

\[(d) \quad f^+ = E_Q(f \mid \mathcal{B}_0) \quad [Q]\]

and

\[(e) \quad f^+ = E_\theta(f \mid \mathcal{B}_0) \quad [P_\theta] \text{ for all } \theta \in \Theta.\]

Then,

\[
\int_X f E_Q(q_\theta \mid \mathcal{B}_0) dQ
\]

\[= \int_X f Q_q \mid \mathcal{B}_0) q_\theta dQ \quad \text{[by the self-adjoint property of conditional expectations]}\]

\[= \int_X f^+ dP_\theta \quad \text{[by (d) and the fact that } q_\theta = dP_\theta / dQ]\]

\[= \int_X f E_Q(q_\theta \mid \mathcal{B}) dQ \quad \text{[by (e) and the fact that } E_Q(q_\theta \mid \mathcal{B}) = dP_\theta / dQ \mid \mathcal{B}]\]

Hence, (c) follows as required.

Let $\mathcal{B}_{\ast \ast} = \sigma(\{E_Q(q_\theta \mid \mathcal{B}) : \theta \in \Theta\})$. Since $E_Q(q_\theta \mid \mathcal{B})$ coincides with the Radon-Nikodym derivative of $P_\theta \mid \mathcal{B}$ with respect to $Q \mid \mathcal{B}$, it is known that $\mathcal{B}_{\ast \ast}$ is a minimum sufficient subfield for $\mathcal{B}$.

Let $\mathcal{B}^\ast = \mathcal{B} \vee \mathcal{B}_{\ast \ast}$. We shall prove that $\mathcal{B}^\ast$ is a minimum adequate subfield for $\mathcal{B}$.

**Theorem D.** Suppose $P$ is dominated by some $\sigma$-finite measure $\lambda$ on $A$. Then $\mathcal{B}^\ast \min \text{adq}(\mathcal{B}; C, P)$. 

25
PROOF. Since $B_{**} \supseteq B^*$ and $B_{**} \subseteq B^*$, it follows that $B^*$ suff $(B, P)$ and $B_{**} \subseteq B^*$. From Lemma B, we know that $B_{**} \subseteq B^*$ with respect to each $P_\theta$, $\theta \in \Theta$. Since $B_{**} \subseteq B^* \subseteq B$, it follows from Proposition D of Section II that $B_{**} \subseteq B^*$ with respect to $P_\theta$ for all $\theta \in \Theta$. Hence, $B^*$ adq $(B; C, P)$.

To prove the minimality of $B^*$, let $B_0$ adq $(B; C, P)$ be given. In view of Lemma C and the minimality of $B_{**}$, we know that $B_0 \subseteq B_{**}$ [P] and $B_{**} \subseteq B_0$ [P].

Therefore, $B^* = B_0 \vee B_{**} \subseteq B_0$ [P].

We now illustrate our main result by the following example.

EXAMPLE. Let $X_1, X_2, \ldots, X_n$ be independent random variables from $N(\theta, 1)$, $-\infty < \theta < \infty$, and let $Y = X_n + Z$, where $Z$ is a random variable from $N(0, 1)$ which is independent of $X = (X_1, X_2, \ldots, X_n)$. The joint density of $X$ and $Y$ is then given by

$$p_\theta(x, y) = \left(\frac{1}{\sqrt{2\pi}}\right)^{n+1} e^{-\frac{1}{2} \sum_{i=1}^{n} (x_i - \theta)^2} - \frac{1}{2} (y - x_n)^2,$$

and the conditional density of $Y$ given $X$ is therefore found to be

$$p_\theta(y|x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} (y - x_n)^2}.$$

Thus, for any $P$-integrable function $S = \mathcal{S}(Y)$.
$E_\theta [g(Y)|X]$ is clearly a function only of $X_n$, namely,

$E_\theta [g(Y)|X] = f(X_n)$.

Let us now take the dominating measure $Q$ as $P_{\theta_0} = \prod_{i=1}^n N(\theta_0, 1)$ for some $\theta_0 \in \Theta$. It is easy to see that

$$q_\theta = \frac{dP_\theta}{dQ} = e^{n(\theta-\theta_0)X - \frac{n}{2}(\theta^2-\theta_0^2)} > 0,$$

and

$$E_Q(q_\theta|X) = e^{n(\theta-\theta_0)\bar{X} - \frac{n}{2}(\theta^2-\theta_0^2)} > 0.$$

Thus, for each $\theta \in \Theta$ and each $g = g(Y)$, $E_\theta = E_\theta [g(Y)|X]$.

Hence,

$$B_* = \sigma\{E_\theta [g(Y)|X] : \theta \in \Theta, g = g(Y) \text{ is } P\text{-integrable}\} = \sigma\{X_n\}.$$

However, it is well known that $\bar{X}$ is minimum sufficient for $X$.

Therefore, we conclude that $(\bar{X}, X_n) \min \text{ad}(\bar{X}; Y, P)$ in accordance with Theorem D.
REFERENCES


