1. *On the concept of contiguity and related theorems*

Summary

The main purpose of this chapter is to present the concept of contiguity (see Definition 2.1) introduced by LeCam [4] and study some alternative characterizations of it (see Theorem 6.1). In the process of doing so, some auxiliary concepts such as weak convergence, relative compactness and tightness of a sequence of probability measures are needed. These concepts are introduced in this chapter, as we go along, and also some of their relationships are stated and/or proved. For the omitted proofs, the reader is always referred to appropriate sources. The various characterizations of continguity provide alternative methods one may employ in establishing the presence (or absence) of contiguity in a given case. Some concrete examples are used for illustrative purposes.

Contiguity is a concept of 'nearness' of sequences of probability measures. It would then be appropriate to relate it to other more familiar concepts of the same nature such as 'nearness' of two sequences of probability measures expressed by the norm (L_1-norm) associated with convergence in variation. By means of examples, it is shown, as one would expect, that 'nearness' of two sequences of probability measures expressed by contiguity is weaker than that expressed by the L_1-norm.

Some attention is also focused to possible relationships between contiguity on the one hand, and mutual absolute continuity and tightness on the other. In connection with this, it is shown, by means of examples, that mutual absolute continuity of the (corresponding) measures in two sequences of probability measures need not imply contiguity of the sequences. Although the converse is not true either, it is always possible to replace two

contiguous sequences of probability measures by two other contiguous sequences whose (corresponding) members are absolutely continuous with respect to one another and these latter sequences lie close to the given ones in the L_1-norm sense (see Theorem 5.1). Also it is shown that tightness need not imply contiguity and contiguous sequences of probability measures need not be tight.

Finally, a number of important theorems (see Theorems 7.1, 7.2 and their corollaries), based on the assumption of contiguity, are formulated and proved. As will be seen in later chapters, these results are very essential for the statistical applications to be discussed in this monograph. Loosely speaking, these results provide the asymptotic distribution of a sequence of statistics, under a given sequence of probability measures $\{P'_n\}$, if the same sequence of statistics has a limiting distribution, under a sequence of probability measures $\{P_n\}$ contiguous to $\{P'_n\}$. For example, $\{P_n\}$ may correspond to a hypothesis being tested and $\{P'_n\}$ to 'close' alternatives.

In order to avoid undue repetition, we should like to mention at the outset that in this chapter, as well as in the subsequent ones, all limits are taken as $\{n\}$, or subsequences thereof, converges to infinity through the positive (or non-negative) integers unless otherwise specified. Also, integrals without limits are understood to be taken over the entire (appropriate) space.

1 Some preliminary definitions and results

In all that follows, $(S, \mathscr{T})$ is a topological space, where the topology $\mathscr{T}$ is defined by a metric on S. $\mathscr{S}$ is the topological Borel σ-field generated by $\mathscr{T}$.

Let $\{\mathscr{L}_n\}$ and $\mathscr{L}$ be a sequence of probability measures and a probability measure, respectively, defined on $\mathscr{S}$. Then

DEFINITION 1.1 We say that $\{\mathscr{L}_n\}$ converges *weakly* to $\mathscr{L}$ and we write $\mathscr{L}_n \Rightarrow \mathscr{L}$ if $\int f \, d\mathscr{L}_n \to \int f \, d\mathscr{L}$ for all real-valued, bounded and continuous functions f defined on S.

If
$$(S, \mathscr{S}) = (R^m, \mathscr{B}^m) = \prod_{j=1}^{m} (R_j, \mathscr{B}_j) \quad (m = 1, 2, \ldots),$$
where $(R_j, \mathscr{B}_j) = (R, \mathscr{B})$, the Borel real line, and if F_n and F are the distribution functions (d.f.s) corresponding to $\mathscr{L}_n$ and $\mathscr{L}$, re-

spectively, then weak convergence is equivalent to the convergence $F_n(x) \to F(x)$, $x \in C(F)$, the set of continuity points of F (this is a generalization of the Helly–Bray theorem discussed, e.g. in Loève [1], p. 182; see also Billingsley [4], p. 18).

Theorems 1.1 A, 1.2 A in the appendix (and also Theorem 1.1 below along with Remark 1.2) provide alternative characterizations of weak convergence.

THEOREM 1.1 Suppose $(S, \mathscr{S}) = (R^m, \mathscr{B}^m)$ and let $\{\mathscr{L}_n\}$ and $\mathscr{L}$ be a sequence of probability measures and a probability measure, respectively, defined on $\mathscr{S}$ such that $\mathscr{L}_n \Rightarrow \mathscr{L}$. Let f be a real-valued, bounded function defined on S such that its restriction to a compact set K is continuous, f vanishes outside K and f need not vanish on the boundary ∂K of K, provided $\mathscr{L}(\partial K) = 0$. Then $\int f \mathrm{d}\mathscr{L}_n \to \int f \mathrm{d}\mathscr{L}$.

Proof Of course, if f is continuous, the conclusion holds true whether or not $\mathscr{L}(\partial K) = 0$. Thus it suffices to restrict ourselves to fs which are discontinuous on ∂K, provided $\mathscr{L}(\partial K) = 0$. Since

$$\int_{K^c} f \mathrm{d}\mathscr{L}_n = \int_{K^c} f \mathrm{d}\mathscr{L} = 0,$$

it suffices to show that

$$\int_K f \mathrm{d}\mathscr{L}_n \to \int_K f \mathrm{d}\mathscr{L}. \tag{1.1}$$

By Theorem 1.1 A, $\mathscr{L}_n(K) \to \mathscr{L}(K)$ since $\mathscr{L}(\partial K) = 0$. Therefore, if $\mathscr{L}(K) = 0$, then

$$\left|\int_K f \mathrm{d}\mathscr{L}_n\right| \leqslant M\mathscr{L}_n(K) \to 0 = \int_K f \mathrm{d}\mathscr{L},$$

where M is an upper bound for $|f|$ on K, and hence (1.1) is true. Thus we assume that $\mathscr{L}(K) > 0$ and on $\mathscr{S} \cap K$, define $\mathscr{L}_n^*$ for all sufficiently large n (so that $\mathscr{L}_n(K) > 0$), and $\mathscr{L}^*$ by

$$\mathscr{L}_n^*(B) = \mathscr{L}_n(B)/\mathscr{L}_n(K), \quad \mathscr{L}^*(B) = \mathscr{L}(B)/\mathscr{L}(K).$$

Let $(\partial B)_r$ be the boundary of a set B with respect to the induced (in K) topology. Then it is clear that $\partial B \subseteq (\partial B)_r \cup \partial K$, so that

1-2

$\mathscr{L}^*[(\partial B)_r] = 0$ implies $\mathscr{L}(\partial B) = 0$. Thus $\mathscr{L}^*$-continuity sets are also $\mathscr{L}$-continuity sets. Therefore $\mathscr{L}_n^* \Rightarrow \mathscr{L}^*$. It follows that

$$\int_K f \,\mathrm{d}\mathscr{L}_n = \mathscr{L}_n(K)\int_K f\,\mathrm{d}\mathscr{L}_n^* \to \mathscr{L}(K)\int_K f\,\mathrm{d}\mathscr{L}^* = \int_K f\,\mathrm{d}\mathscr{L}.$$

This establishes (1.1) and hence the theorem itself. ∎

COROLLARY 1.1 Let f be as in the theorem except that f is equal to a constant c on K^c and f need not be equal to c on ∂K, provided $\mathscr{L}(\partial K) = 0$. Then $\int f\,\mathrm{d}\mathscr{L}_n \to \int f\,\mathrm{d}\mathscr{L}$.

Proof Replace f by $f-c$ and apply the theorem. ∎

REMARK 1.1 By Theorem 1.2A, it follows that the converse of Theorem 1.1 is also true. That is, if $\int f\,\mathrm{d}\mathscr{L}_n \to \int f\,\mathrm{d}\mathscr{L}$ for each f as described in Theorem 1.1, then $\mathscr{L}_n \Rightarrow \mathscr{L}$.

We recall that $\{\mathscr{L}_n\}$ is a sequence of probability measures defined on $\mathscr{S}$. Then

DEFINITION 1.2 The sequence $\{\mathscr{L}_n\}$ is said to be *relatively compact* if for every subsequence $\{n'\} \subseteq \{n\}$ there exists a further subsequence $\{n''\} \subseteq \{n'\}$ such that $\{\mathscr{L}_{n''}\}$ converges weakly to a probability measure (which, in general, depends on $\{n''\}$).

The following definition will also be useful.

DEFINITION 1.3 The sequence $\{\mathscr{L}_n\}$ is said to be *tight* if for every $\epsilon > 0$ there is a compact set $K = K(\epsilon)$ such that

$$\mathscr{L}_n(K) > 1-\epsilon$$

for all n.

If $(S, \mathscr{S}) = (R^m, \mathscr{B}^m)$, this condition can be replaced by:

$$\mathscr{L}_n([a,b]) > 1-\epsilon$$

for all n, where $[a,b]$ is a closed interval in R^m. This is so because any compact set in R^m can be enclosed in a bounded, closed interval, and bounded, closed intervals are compact.

The concepts of relative compactness and tightness are related as follows.

THEOREM 1.2 Let $(S, \mathscr{S}) = (R^m, \mathscr{B}^m)$. Then $\{\mathscr{L}_n\}$ is relatively compact if and only if it is tight.

By Theorem 1.3A, tightness always implies relative compactness. The reverse implication is also true, by the same theorem, provided S is separable and complete. This requirement is, clearly, satisfied in the present situation, where, of course, $\mathscr{T}$ is assumed to be the usual topology in R^m. So Theorem 1.2 is a special case of Theorem 1.3A. However, a simple proof of Theorem 1.2 can be presented along the following lines.

Proof of Theorem 1.2 Suppose that $\{\mathscr{L}_n\}$ is tight. Then, for every $\epsilon > 0$, there exists a closed interval in R^m, $[a, b]$, depending on ϵ, such that $\mathscr{L}_n([a,b]) > 1-\epsilon$ for all n. Let F_n be the d.f. corresponding to $\mathscr{L}_n$. Then by the weak compactness theorem for d.f.s, for every $\{n'\} \subset \{n\}$ there exists $\{n''\} \subset \{n'\}$ such that $F_{n''}(x) \to F^*(x)$ for all $x \in C(F^*)$, where F^* is a d.f. in all other respects except that it might fail to have variation equal to one. We shall, actually, show that this does not occur. To this end, let $\mathscr{L}^*$ be the measure induced by F^*. Then $\mathscr{L}^*$ is a probability measure. In fact, $\mathscr{L}_{n''}([a,b]) > 1-\epsilon$ for all n'' and a, b can be taken to be continuity points of F^*. Taking the limit as $n'' \to \infty$, we have $\mathscr{L}^*([a,b]) \geqslant 1-\epsilon$. Thus, for every $\epsilon > 0$, there exists a closed interval $[a, b]$ in R^m, depending on ϵ, such that

$$\mathscr{L}^*([a,b]) \geqslant 1-\epsilon,$$

and this implies that $\mathscr{L}^*$ is a probability measure. Hence, tightness implies relative compactness.

Now assume that $\{\mathscr{L}_n\}$ is relatively compact. Then, for every $\epsilon > 0$, we claim that there exists a closed interval $[a, b]$ in R^m, depending on ϵ, such that $\mathscr{L}_n([a,b]) > 1-\epsilon$ for all n. In fact, if this were not true, then for some $\epsilon > 0$ and for every $[a, b]$ in R^m there would exist a k, depending on ϵ and $[a, b]$, such that $\mathscr{L}_k([a,b]) \leqslant 1-\epsilon$. Apply this argument for $a_n, b_n \in R^m$, where $a_n = (-n, \ldots, -n)'$, $b_n = (n, \ldots, n)'$ and '$'$' denotes transpose. Thus, for every n, there exists n' such that $\mathscr{L}_{n'}([a_n, b_n]) \leqslant 1-\epsilon$, where $\{n'\} \subseteq \{n\}$ and $n' \to \infty$. Then there cannot exist $\{n''\} \subseteq \{n'\}$ such that $\{\mathscr{L}_{n''}\}$ converges (weakly) to a probability measure $\mathscr{L}$, say. This is so because, if F is the d.f. corresponding to $\mathscr{L}$, then for every $x, y \in C(F)$ with $x < y$ (to be understood in the coordinatewise sense), we have

$$\mathscr{L}((x,y]) = \lim \mathscr{L}_{n''}((x,y]) \leqslant \liminf \mathscr{L}_{n''}((a_{n^*}, b_{n^*}]) \leqslant 1-\epsilon,$$

where n^* is the subscript of the interval associated with n'' and the limits are taken as $n^* \to \infty$ which implies $n'' \to \infty$. Thus $\mathscr{L}((x, y]) \leqslant 1-\epsilon$ for every $x, y \in C(F)$, which implies that $\mathscr{L}$ is not a probability measure. This contradicts our assumption of $\{\mathscr{L}_n\}$ being relatively compact. The proof of the theorem is completed. ∎

Now let $\{(\mathscr{X}, \mathscr{A}_n, P_n)\}$ be a sequence of probability spaces and let $\{T_n\}$ be a sequence of m-dimensional random vectors such that T_n is $\mathscr{A}_n$-measurable. Set $\mathscr{L}_n = \mathscr{L}(T_n|P_n)$. Then the following proposition will prove useful on many occasions.

PROPOSITION 1.1 The sequence $\{\mathscr{L}_n\}$ just defined is relatively compact (equivalently, tight) if and only if, for every $\epsilon > 0$, there exists $b = b(\epsilon) > 0$ such that $P_n(\|T_n\| > b) < \epsilon$ for all n, where $\|\cdot\|$ denotes the usual Euclidean norm in R^m.

Proof By Theorem 1.2, $\{\mathscr{L}_n\}$ is relatively compact if and only if it is tight. On the other hand, tightness of $\{\mathscr{L}_n\}$ is equivalent to the existence of an interval $[a, b]$ in R^m, depending on ϵ, such that $\mathscr{L}_n([a, b]) > 1-\epsilon$. This is so by the comments following Definition 1.3. But $\mathscr{L}_n([a, b]) = P_n(T_n \in [a, b])$. Thus $\{\mathscr{L}_n\}$ is tight if and only if $P_n(T_n \in [a, b]) > 1-\epsilon$ for all n. This, however, is equivalent to the existence of a $c(\epsilon) = c = (c_1, \ldots, c_m)'$ with positive coordinates such that $P_n(T_n \in [-c, c]) > 1-\epsilon$ for all n. Taking $b(\epsilon) = \|c(\epsilon)\|$, we get the desired result. ∎

Let now P and Q be two probability measures on the σ-field $\mathscr{A}$ of subsets of $\mathscr{X}$. Then the L_1-*norm* of $P-Q$, denoted by $\|P-Q\|$, is defined by $\|P-Q\| = 2\sup\{|P(A)-Q(A)|; A \in \mathscr{A}\}$. Next, if f and g are densities of P and Q, respectively, relative to a dominating σ-finite measure μ (e.g. $\mu = P+Q$), then

$$\|P-Q\| = \int |f-g| \, d\mu.$$

This is shown in Theorem 1.4A.

The following result is isolated here for convenient reference.

PROPOSITION 1.2 For each n, let P_n and Q_n be probability measures on $\mathscr{A}_n$ and let $\{T_n\}$ be a sequence of m-dimensional random vectors such that T_n is $\mathscr{A}_n$-measurable. Set

$$\mathscr{L}_n^P = \mathscr{L}(T_n|P_n) \quad \text{and} \quad \mathscr{L}_n^Q = \mathscr{L}(T_n|Q_n),$$

and suppose that $\|P_n - Q_n\| \to 0$. Then we have

(i) $$\|\mathscr{L}_n^P - \mathscr{L}_n^Q\| \to 0.$$

(ii) If $\mathscr{L}_n^P \Rightarrow \mathscr{L}$, a probability measure, then $\mathscr{L}_n^Q \Rightarrow \mathscr{L}$ and conversely.

Proof (i) For $B \in \mathscr{B}^m$, one has

$$|\mathscr{L}_n^P(B) - \mathscr{L}_n^Q(B)| = |P_n(T_n \in B) - Q_n(T_n \in B)|,$$

so that

$$\|\mathscr{L}_n^P - \mathscr{L}_n^Q\| = 2\sup\,[|\mathscr{L}_n^P(B) - \mathscr{L}_n^Q(B)|;\, B \in \mathscr{B}^m]$$
$$\leqslant 2\sup\,[|P_n(A) - Q_n(A)|;\, A \in \mathscr{A}_n] = \|P_n - Q_n\| \to 0.$$

(ii) Let B be a continuity set of $\mathscr{L}$. Then we have

$$|\mathscr{L}_n^Q(B) - \mathscr{L}(B)| \leqslant |\mathscr{L}_n^Q(B) - \mathscr{L}_n^P(B)| + |\mathscr{L}_n^P(B) - \mathscr{L}(B)|.$$

If $\mathscr{L}_n^P \Rightarrow \mathscr{L}$, then the second term on the right-hand side of the above inequality tends to zero, by Theorem 1.1 A. The first term also converges to zero by (i). Therefore, $\mathscr{L}_n^Q \Rightarrow \mathscr{L}$. That $\mathscr{L}_n^Q \Rightarrow \mathscr{L}$ implies $\mathscr{L}_n^P \Rightarrow \mathscr{L}$ follows by symmetry. The proof of the proposition is concluded. ▮

2 Contiguity and its relation to other concepts of 'nearness' of sequences of probability measures

In this section, the concept of contiguity is introduced and some alternative characterizations of it are studied. Its relationship to other familiar concepts of 'nearness' of sequences of probability measures is investigated, and, finally, a number of illustrative examples are discussed.

Let $\{(\mathscr{X}, \mathscr{A}_n)\}$ be a sequence of measurable spaces, and let P_n, P_n' be probability measures on $\mathscr{A}_n$. Also, let $\{T_n\}$ be a sequence of random variables (r.v.s) such that T_n is $\mathscr{A}_n$-measurable.

Definition 2.1 The sequences of probability measures $\{P_n\}$ and $\{P_n'\}$ are said to be *contiguous* if the following is true: for any $\mathscr{A}_n$-measurable r.v.s T_n, $T_n \to 0$ in P_n-probability if and only if $T_n \to 0$ in P_n'-probability.

The concept of contiguity is a concept expressing 'closeness' or 'nearness' between the sequences of probability measures $\{P_n\}$ and $\{P'_n\}$ in the sense of the definition just given. Some more light is shed on this concept by the fact that, if $A_n \in \mathscr{A}_n$, then $P_n(A_n) \to 0$ if and only if $P'_n(A_n) \to 0$, provided $\{P_n\}$ and $\{P'_n\}$ are contiguous, as will be shown below. Also some concrete examples will further illustrate the point.

REMARK 2.1 It is worth noticing that contiguity is transitive. That is, if $\{P_n\}$, $\{Q_n\}$ and $\{Q_n\}$, $\{R_n\}$ are contiguous, then $\{P_n\}$, $\{R_n\}$ are contiguous. This is an immediate consequence of Definition 2.1.

PROPOSITION 2.1 The sequences of probability measures $\{P_n\}$ and $\{P'_n\}$ are contiguous if and only if, for $A_n \in \mathscr{A}_n$, $P_n(A_n) \to 0$ if and only if $P'_n(A_n) \to 0$.

Proof Assume $\{P_n\}$ and $\{P'_n\}$ to be contiguous, and let $P_n(A_n) \to 0$ with $A_n \in \mathscr{A}_n$. Set $T_n = I_{A_n}$. Then, for every

$$(1 >)\,\epsilon > 0, \quad P_n(|T_n| > \epsilon) = P_n(A_n),$$

so that $T_n \to 0$ in P_n-probability. This implies $T_n \to 0$ in P'_n-probability by Definition 2.1. Since $P'_n(|T_n| > \epsilon) = P'_n(A_n)$, we have then $P'_n(A_n) \to 0$. The fact that $P'_n(A_n) \to 0$ implies $P_n(A_n) \to 0$ is treated entirely symmetrically. For the converse, we have: let $T_n \to 0$ in P_n-probability and, for $\epsilon > 0$, set

$$A_n = (|T_n| > \epsilon).$$

Then $P_n(A_n) \to 0$ and this implies $P'_n(A_n) \to 0$. However, this last convergence is equivalent to the convergence $T_n \to 0$ in P'_n-probability. That $T_n \to 0$ in P'_n-probability implies $T_n \to 0$ in P_n-probability follows by symmetry. ∎

Of course, if $\|P_n - P'_n\| \to 0$, then the sequences $\{P_n\}$ and $\{P'_n\}$ are as close together as they can be, apart from having identical elements. One would then expect that convergence in L_1-norm would imply contiguity. That this is, actually, the case is seen in the following lemma.

LEMMA 2.1 If $\|P_n - P'_n\| \to 0$, then $\{P_n\}$ and $\{P'_n\}$ are contiguous.

Proof The convergence $\|P_n - P'_n\| \to 0$ implies that, for every $\epsilon > 0$, $|P_n(A_n) - P'_n(A_n)| < \epsilon$ for every $A_n \in \mathscr{A}_n$ and all sufficiently large n, $n \geqslant n_1$, say. Thus $P_n(A_n) < \epsilon$ implies $P'_n(A_n) < 2\epsilon$ and $P'_n(A_n) < \epsilon$ implies $P_n(A_n) < 2\epsilon$ for every $A_n \in \mathscr{A}_n$ and all $n \geqslant n_1$. Then Proposition 2.1 applies and gives the desired result. ∎

REMARK 2.2 It is demonstrated by means of examples (see, e.g. Example 3.1 (i)) that the converse of Lemma 2.1 is not true. Thus contiguity is a weaker 'measure of nearness' of sequences of probability measures than that expressed by convergence in L_1-norm.

We now consider the following example as a simple application of Lemma 2.1.

EXAMPLE 2.1 Let $(\mathscr{X}, \mathscr{A}_n) = (R, \mathscr{B})$, $P_n = U(-1/n, 1)$, the uniform measure over $(-1/n, 1)$ and $P'_n = U(0, 1+1/n)$, the uniform measure over $(0, 1+1/n)$.

Then, if $$f_n = \mathrm{d}P_n/\mathrm{d}l \quad \text{and} \quad g_n = \mathrm{d}P'_n/\mathrm{d}l,$$
where l is the Lebesgue measure in R, we have
$$\|P_n - P'_n\| = \int |f_n - g_n| \, \mathrm{d}l = 2/(n+1) \to 0.$$
Thus $\{P_n\}$ and $\{P'_n\}$ are contiguous, by Lemma 2.1.

REMARK 2.3 Example 2.1 also shows that contiguity need not imply mutual absolute continuity (either for all n or only for all sufficiently large n), as one might be tempted to (wrongly) infer by Proposition 2.1. It will be shown later, however, that any given pair of contiguous sequences of probability measures can always be replaced by another pair of contiguous sequences whose members are mutually absolutely continuous and lie arbitrarily close to the given ones (for this, see Theorem 5.1). On the other hand, mutual absolute continuity need not imply contiguity either, as is demonstrated by the following example.

EXAMPLE 2.2 Let
$$(\mathscr{X}, \mathscr{A}_n) = (R, \mathscr{B}), \quad P_n = N(\mu_n, 1) \quad \text{and} \quad P'_n = N(\mu'_n, 1),$$
where $\mu_n \to -\infty$ and $\mu'_n \to \infty$.

Then, clearly, $P_n \approx P'_n$ for all n. (As usual, by $P \approx Q$ we express the fact that $P \ll Q$ and $Q \ll P$.) Consider the set A_n defined by

$A_n = (\mu_n - 1, \mu_n + 1)$. Then $P_n(A_n) = c$, a positive constant ($c \approx 0{\cdot}68$), so that $P_n(A_n)$ does not tend to zero. But $P'_n(A_n) \to 0$, clearly. Thus, by Proposition 2.1, $\{P_n\}$ and $\{P'_n\}$ cannot be contiguous.

REMARK 2.4 From this example, we might (wrongly) conclude that the absence of contiguity of $\{P_n\}$, $\{P'_n\}$ is due to the lack of tightness. (It is clear that the sequences are not tight.) That tightness is not a necessary condition for contiguity is shown in Example 3.1 (i), where the sequences involved are not tight (for a special choice of the parameters) and yet contiguous. In the same example and for another choice of the parameters, it is also shown that tightness is not a sufficient condition for contiguity.

Summarizing some of the results obtained so far, we have:

If $\|P_n - P'_n\| \to 0$, then $\{P_n\}$, $\{P'_n\}$ are contiguous. The converse need not be true.

Contiguity of $\{P_n\}$, $\{P'_n\}$ need not imply $P_n \approx P'_n$ for all or sufficiently large n (see, however, Theorem 5.1). The converse is also true.

Contiguity of $\{P_n\}$, $\{P'_n\}$ need not imply their tightness. The converse is also true.

The following remark is relevant here.

REMARK 2.5 If $\{P_n\}$ and $\{P'_n\}$ are contiguous, but $P_n \not\approx P'_n$, then it follows from Proposition 2.1 that the support of the singular part of P_n with respect to P'_n has P_n-probability tending to zero, and the support of the singular part of P'_n with respect to P_n has P'_n-probability tending to zero.

3 Alternative characterizations of contiguity

Proposition 2.1 provides an alternative characterization of contiguity. Additional ones will be given in the sequel but for their formulation and proof, we need some more notation and also some auxiliary results. To this end, for each n, let P_n, P'_n be probability measures on $\mathscr{A}_n$ and let μ_n be a σ-finite measure dominating them. Furthermore, let

$$f_n = \mathrm{d}P_n/\mathrm{d}\mu_n, \quad g_n = \mathrm{d}P'_n/\mathrm{d}\mu_n. \tag{3.1}$$