

### 3.13: Absolute Continuity and Density Functions

#### Basic Theory

Our starting point is a measurable space  $(S, \mathcal{S})$ . That is  $S$  is a set and  $\mathcal{S}$  is a  $\sigma$ -algebra of subsets of  $S$ . In the last section, we discussed general measures on  $(S, \mathcal{S})$  that can take positive and negative values. Special cases are positive measures, finite measures, and our favorite kind, probability measures. In particular, we studied properties of general measures, ways to construct them, special sets (positive, negative, and null), and the Hahn and Jordan decompositions.

In this section, we see how to construct a new measure from a given positive measure using a density function, and we answer the fundamental question of when a measure has a density function relative to the given positive measure.

#### Relations on Measures

The answer to the question involves two important relations on the collection of measures on  $(S, \mathcal{S})$  that are defined in terms of null sets. Recall that  $A \in \mathcal{S}$  is *null* for a measure  $\mu$  on  $(S, \mathcal{S})$  if  $\mu(B) = 0$  for every  $B \in \mathcal{S}$  with  $B \subseteq A$ . At the other extreme,  $A \in \mathcal{S}$  is a *support set* for  $\mu$  if  $A^c$  is a null set. Here are the basic definitions:

Suppose that  $\mu$  and  $\nu$  are measures on  $(S, \mathcal{S})$ .

1.  $\nu$  is *absolutely continuous* with respect to  $\mu$  if every null set of  $\mu$  is also a null set of  $\nu$ . We write  $\nu \ll \mu$ .
2.  $\mu$  and  $\nu$  are *mutually singular* if there exists  $A \in \mathcal{S}$  such that  $A$  is null for  $\mu$  and  $A^c$  is null for  $\nu$ . We write  $\mu \perp \nu$ .

Thus  $\nu \ll \mu$  if every support set of  $\mu$  is a support set of  $\nu$ . At the opposite end,  $\mu \perp \nu$  if  $\mu$  and  $\nu$  have disjoint support sets.

Suppose that  $\mu, \nu$ , and  $\rho$  are measures on  $(S, \mathcal{S})$ . Then

1.  $\mu \ll \mu$ , the *reflexive property*.
2. If  $\mu \ll \nu$  and  $\nu \ll \rho$  then  $\mu \ll \rho$ , the *transitive property*.

Recall that every relation that is reflexive and transitive leads to an equivalence relation, and then in turn, the original relation can be extended to a partial order on the collection of equivalence classes. This general theorem on relations leads to the following two results.

Measures  $\mu$  and  $\nu$  on  $(S, \mathcal{S})$  are *equivalent* if  $\mu \ll \nu$  and  $\nu \ll \mu$ , and we write  $\mu \equiv \nu$ . The relation  $\equiv$  is an equivalence relation on the collection of measures on  $(S, \mathcal{S})$ . That is, if  $\mu, \nu$ , and  $\rho$  are measures on  $(S, \mathcal{S})$  then

1.  $\mu \equiv \mu$ , the *reflexive property*
2. If  $\mu \equiv \nu$  then  $\nu \equiv \mu$ , the *symmetric property*
3. If  $\mu \equiv \nu$  and  $\nu \equiv \rho$  then  $\mu \equiv \rho$ , the *transitive property*

Thus,  $\mu$  and  $\nu$  are equivalent if they have the same null sets and thus the same support sets. This equivalence relation is rather weak: equivalent measures have the same support sets, but the values assigned to these sets can be very different. As usual, we will write  $[\mu]$  for the equivalence class of a measure  $\mu$  on  $(S, \mathcal{S})$ , under the equivalence relation  $\equiv$ .

If  $\mu$  and  $\nu$  are measures on  $(S, \mathcal{S})$ , we write  $[\mu] \preceq [\nu]$  if  $\mu \ll \nu$ . The definition is consistent, and defines a partial order on the collection of equivalence classes. That is, if  $\mu, \nu$ , and  $\rho$  are measures on  $(S, \mathcal{S})$  then

1.  $[\mu] \preceq [\mu]$ , the *reflexive property*.
2. If  $[\mu] \preceq [\nu]$  and  $[\nu] \preceq [\mu]$  then  $[\mu] = [\nu]$ , the *antisymmetric property*.
3. If  $[\mu] \preceq [\nu]$  and  $[\nu] \preceq [\rho]$  then  $[\mu] \preceq [\rho]$ , the *transitive property*

The singularity relation is trivially symmetric and is almost anti-reflexive.

Suppose that  $\mu$  and  $\nu$  are measures on  $(S, \mathcal{S})$ . Then

1. If  $\mu \perp \nu$  then  $\nu \perp \mu$ , the *symmetric property*.
2.  $\mu \perp \mu$  if and only if  $\mu = \mathbf{0}$ , the *zero measure*.

**Proof**

Part (a) is trivial from the symmetry of the definition. For part (b), note that  $S$  is null for  $\mathbf{0}$  and  $\emptyset$  is null for  $\mathbf{0}$ , so  $\mathbf{0} \perp \mathbf{0}$ . Conversely, suppose that  $\mu$  is a measure and  $\mu \perp \mu$ . Then there exists  $A \in \mathcal{S}$  such that  $A$  is null for  $\mu$  and  $A^c$  is null for  $\mu$ . But then  $S = A \cup A^c$  is null for  $\mu$ , so  $\mu(B) = 0$  for every  $B \in \mathcal{S}$ .

Absolute continuity and singularity are preserved under multiplication by nonzero constants.

Suppose that  $\mu$  and  $\nu$  are measures on  $(S, \mathcal{S})$  and that  $a, b \in \mathbb{R} \setminus \{0\}$ . Then

1.  $\nu \ll \mu$  if and only if  $a\nu \ll b\mu$ .
2.  $\nu \perp \mu$  if and only if  $a\nu \perp b\mu$ .

Proof

Recall that if  $c \neq 0$ , then  $A \in \mathcal{S}$  is null for  $\mu$  if and only if  $A$  is null for  $c\mu$ .

There is a corresponding result for sums of measures.

Suppose that  $\mu$  is a measure on  $(S, \mathcal{S})$  and that  $\nu_i$  is a measure on  $(S, \mathcal{S})$  for each  $i$  in a countable index set  $I$ . Suppose also that  $\nu = \sum_{i \in I} \nu_i$  is a well-defined measure on  $(S, \mathcal{S})$ .

1. If  $\nu_i \ll \mu$  for every  $i \in I$  then  $\nu \ll \mu$ .
2. If  $\nu_i \perp \mu$  for every  $i \in I$  then  $\nu \perp \mu$ .

Proof

Recall that if  $A \in \mathcal{S}$  is null for  $\nu_i$  for each  $i \in I$ , then  $A$  is null for  $\nu = \sum_{i \in I} \nu_i$ , assuming that this is a well-defined measure.

As before, note that  $\nu = \sum_{i \in I} \nu_i$  is well-defined if  $\nu_i$  is a positive measure for each  $i \in I$  or if  $I$  is finite and  $\nu_i$  is a finite measure for each  $i \in I$ . We close this subsection with a couple of results that involve both the absolute continuity relation and the singularity relation

Suppose that  $\mu, \nu$ , and  $\rho$  are measures on  $(S, \mathcal{S})$ . If  $\nu \ll \mu$  and  $\mu \perp \rho$  then  $\nu \perp \rho$ .

Proof

Since  $\mu \perp \rho$ , there exists  $A \in \mathcal{S}$  such that  $A$  is null for  $\mu$  and  $A^c$  is null for  $\rho$ . But  $\nu \ll \mu$  so  $A$  is null for  $\nu$ . Hence  $\nu \perp \rho$ .

Suppose that  $\mu$  and  $\nu$  are measures on  $(S, \mathcal{S})$ . If  $\nu \ll \mu$  and  $\nu \perp \mu$  then  $\nu = \mathbf{0}$ .

Proof

From the [previous theorem](#) (with  $\rho = \nu$ ) we have  $\nu \perp \nu$  and hence by (5),  $\nu = \mathbf{0}$ .

## Density Functions

We are now ready for our study of density functions. Throughout this subsection, we assume that  $\mu$  is a positive,  $\sigma$ -finite measure on our measurable space  $(S, \mathcal{S})$ . Recall that if  $f: S \rightarrow \mathbb{R}$  is measurable, then the integral of  $f$  with respect to  $\mu$  may exist as a number in  $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, \infty\}$  or may fail to exist.

Suppose that  $f: S \rightarrow \mathbb{R}$  is a measurable function whose integral with respect to  $\mu$  exists. Then function  $\nu$  defined by

$$\nu(A) = \int_A f d\mu, \quad A \in \mathcal{S} \quad (3.13.1)$$

is a  $\sigma$ -finite measure on  $(S, \mathcal{S})$  that is absolutely continuous with respect to  $\mu$ . The function  $f$  is a *density function* of  $\nu$  relative to  $\mu$ .

Proof

To say that the integral exists means that either  $\int_S f^+ d\mu < \infty$  or  $\int_S f^- d\mu < \infty$ , where as usual,  $f^+$  and  $f^-$  are the positive and negative parts of  $f$ . So  $\nu(A) = \nu_+(A) - \nu_-(A)$  for  $A \in \mathcal{S}$  where  $\nu_+(A) = \int_A f^+ d\mu$  and  $\nu_-(A) = \int_A f^- d\mu$ . Both  $\nu_+$  and  $\nu_-$  are positive measures by basic properties of the integral: Generically, suppose  $g: S \rightarrow [0, \infty)$  is measurable. The integral over the empty set is always 0, so  $\int_\emptyset g d\mu = 0$ . Next, if  $\{A_i : i \in I\}$  is a countable, disjoint collection of sets in  $\mathcal{S}$  and  $A = \bigcup_{i \in I} A_i$ , then by the additivity property of the integral over disjoint domains,

$$\int_A g d\mu = \sum_{i \in I} \int_{A_i} g d\mu \quad (3.13.2)$$

By the assumption that the integral exists, either  $\nu_+$  or  $\nu_-$  is a finite positive measure, and hence  $\nu$  is a measure. As you might guess,  $\nu_+$  and  $\nu_-$  form the Jordan decomposition of  $\nu$ , a point that we will revisit below.

Again, either  $\nu_+$  or  $\nu_-$  is a finite measure. By symmetry, let's suppose that  $\nu_-$  is finite. Then to show that  $\nu$  is  $\sigma$ -finite, we just need to show that  $\nu_+$  is  $\sigma$ -finite. Since  $\mu$  has this property, there exists a collection  $\{A_n : n \in \mathbb{N}_+\}$  with  $A_n \in \mathcal{S}$ ,  $\mu(A_n) < \infty$ , and  $\bigcup_{n=1}^\infty A_n = S$ . Let  $B_n = \{x \in S : f^+(x) \leq n\}$  for  $n \in \mathbb{N}_+$ . Then  $B_n \in \mathcal{S}$  for  $n \in \mathbb{N}_+$  and  $\bigcup_{n=1}^\infty B_n = S$ . Hence  $\{A_m \cap B_n : (m, n) \in \mathbb{N}_+^2\}$  is a countable collection of measurable sets whose union is also  $S$ . Moreover,

$$\nu_+(A_m \cap B_n) = \int_{A_m \cap B_n} f^+ d\mu \leq n\mu(A_m \cap B_n) < \infty \quad (3.13.3)$$

Finally, suppose  $A \in \mathcal{S}$  is a null set of  $\mu$ . If  $B \in \mathcal{S}$  and  $B \subseteq A$  then  $\mu(B) = 0$  so  $\nu(B) = \int_B f d\mu = 0$ . Hence  $\nu \ll \mu$ .

The following three special cases are the most important:

1. If  $f$  is nonnegative (so that the integral exists in  $\mathbb{R} \cup \{\infty\}$ ) then  $\nu$  is a *positive measure* since  $\nu(A) \geq 0$  for  $A \in \mathcal{S}$ .
2. If  $f$  is integrable (so that the integral exists in  $\mathbb{R}$ ), then  $\nu$  is a *finite measure* since  $\nu(A) \in \mathbb{R}$  for  $A \in \mathcal{S}$ .
3. If  $f$  is nonnegative and  $\int_S f d\mu = 1$  then  $\nu$  is a *probability measure* since  $\nu(A) \geq 0$  for  $A \in \mathcal{S}$  and  $\nu(S) = 1$ .

In case 3,  $f$  is the *probability density function* of  $\nu$  relative to  $\mu$ , our favorite kind of density function. When they exist, density functions are essentially unique.

Suppose that  $\nu$  is a  $\sigma$ -finite measure on  $(S, \mathcal{S})$  and that  $\nu$  has density function  $f$  with respect to  $\mu$ . Then  $g: S \rightarrow \mathbb{R}$  is a density function of  $\nu$  with respect to  $\mu$  if and only if  $f = g$  almost everywhere on  $S$  with respect to  $\mu$ .

**Proof**

These results also follow from basic properties of the integral. Suppose that  $f, g: S \rightarrow \mathbb{R}$  are measurable functions whose integrals with respect to  $\mu$  exist. If  $g = f$  almost everywhere on  $S$  with respect to  $\mu$  then  $\int_A f d\mu = \int_A g d\mu$  for every  $A \in \mathcal{S}$ . Hence if  $f$  is a density function for  $\nu$  with respect to  $\mu$  then so is  $g$ . For the converse, if  $\int_A f d\mu = \int_A g d\mu$  for every  $A \in \mathcal{S}$ , then since  $\mu$  is  $\sigma$ -finite, it follows that  $f = g$  almost everywhere on  $S$  with respect to  $\mu$ .

The essential uniqueness of density functions can fail if the positive measure space  $(S, \mathcal{S}, \mu)$  is not  $\sigma$ -finite. A simple example is given [below](#). Our next result answers the question of when a measure has a density function with respect to  $\mu$ , and is the fundamental theorem of this section. The theorem is in two parts: Part (a) is the *Lebesgue decomposition theorem*, named for our old friend Henri Lebesgue. Part (b) is the *Radon-Nikodym theorem*, named for Johann Radon and Otto Nikodym. We combine the theorems because our proofs of the two results are inextricably linked.

Suppose that  $\nu$  is a  $\sigma$ -finite measure on  $(S, \mathcal{S})$ .

1. **Lebesgue Decomposition Theorem.**  $\nu$  can be uniquely decomposed as  $\nu = \nu_c + \nu_s$  where  $\nu_c \ll \mu$  and  $\nu_s \perp \mu$ .
2. **Radon-Nikodym Theorem.**  $\nu_c$  has a density function with respect to  $\mu$ .

**Proof**

The proof proceeds in stages. we first prove the result for finite, positive measures, then for  $\sigma$ -finite, positive measures, and finally for general  $\sigma$ -finite measures. The first stage is the most complicated.

**Part 1**, suppose that  $\mu$  and  $\nu$  are positive, finite measures. Let  $\mathcal{F}$  denote the collection of measurable functions  $g: S \rightarrow [0, \infty)$  with  $\int_A g d\mu \leq \nu(A)$  for all  $A \in \mathcal{S}$ . Note that  $\mathcal{F} \neq \emptyset$  since the constant function 0 is in  $\mathcal{F}$ . The proof works by finding a maximal element of  $\mathcal{F}$  and using this function as the density function of the absolutely continuous part of  $\nu$ .

Our first step is to show that  $\mathcal{F}$  is closed under the max operator. Let  $g_1, g_2 \in \mathcal{F}$ . For  $A \in \mathcal{S}$ , let  $A_1 = \{x \in A : g_1(x) \geq g_2(x)\}$  and  $A_2 = \{x \in A : g_1(x) < g_2(x)\}$ . Then  $A_1, A_2 \in \mathcal{S}$  partition  $A$  so

$$\int_A \max\{g_1, g_2\} d\mu = \int_{A_1} \max\{g_1, g_2\} d\mu + \int_{A_2} \max\{g_1, g_2\} d\mu = \int_{A_1} g_1 d\mu + \int_{A_2} g_2 d\mu \leq \nu(A_1) + \nu(A_2) = \nu(A) \quad (3.13.4)$$

Hence  $\max\{g_1, g_2\} \in \mathcal{F}$ .

Our next step is to show that  $\mathcal{F}$  is closed with respect to increasing limits. Thus suppose that  $g_n \in \mathcal{F}$  for  $n \in \mathbb{N}_+$  and that  $g_n$  is increasing in  $n$  on  $S$ . Let  $g = \lim_{n \rightarrow \infty} g_n$ . Then  $g: S \rightarrow [0, \infty]$  is measurable, and by the monotone convergence theorem,  $\int_A g d\mu = \lim_{n \rightarrow \infty} \int_A g_n d\mu$  for every  $A \in \mathcal{S}$ . But  $\int_A g_n d\mu \leq \nu(A)$  for every  $n \in \mathbb{N}_+$  so  $\int_A g d\mu \leq \nu(A)$ . In particular,  $\int_S g d\mu \leq \nu(S) < \infty$  so  $g < \infty$  almost everywhere on  $S$  with respect to  $\mu$ . Thus, by redefining  $g$  on a  $\mu$ -null set if necessary, we can assume  $g < \infty$  on  $S$ . Hence  $g \in \mathcal{F}$ .

Now let  $\alpha = \sup \left\{ \int_S g d\mu : g \in \mathcal{F} \right\}$ . Note that  $\alpha \leq \nu(S) < \infty$ . By definition of the supremum, for each  $n \in \mathbb{N}_+$  there exist  $g_n \in \mathcal{F}$  such that  $\int_S g_n d\mu > \alpha - \frac{1}{n}$ . Now let  $f_n = \max\{g_1, g_2, \dots, g_n\}$  for  $n \in \mathbb{N}_+$ . Then  $f_n \in \mathcal{F}$  and  $f_n$  is increasing in  $n \in \mathbb{N}_+$  on  $S$ . Hence  $f = \lim_{n \rightarrow \infty} f_n \in \mathcal{F}$  and  $\int_S f d\mu = \lim_{n \rightarrow \infty} \int_S f_n d\mu$ . But  $\int_S f_n d\mu \geq \int_S g_n d\mu > \alpha - \frac{1}{n}$  for each  $n \in \mathbb{N}_+$  and hence  $\int_S f d\mu \geq \alpha$ .

Define  $\nu_c(A) = \int_A f d\mu$  and  $\nu_s(A) = \nu(A) - \nu_c(A)$  for  $A \in \mathcal{S}$ . Then  $\nu_c$  and  $\nu_s$  are finite, positive measures and by our previous theorem,  $\nu_c$  is absolutely continuous with respect to  $\mu$  and has density function  $f$ . Our next step is to show that  $\nu_s$  is singular with respect to  $\mu$ . For  $n \in \mathbb{N}$ , let  $(P_n, P_n^c)$  denote a Hahn decomposition of the measure  $\nu_s - \frac{1}{n}\mu$ . Then

$$\int_A \left( f + \frac{1}{n} \mathbf{1}_{P_n} \right) d\mu = \nu_c(A) + \frac{1}{n} \mu(P_n \cap A) = \nu(A) - \left[ \nu_s(A) - \frac{1}{n} \mu(P_n \cap A) \right] \quad (3.13.5)$$

But  $\nu_s(A) - \frac{1}{n}\mu(P_n \cap A) \geq \nu_s(A \cap P_n) - \frac{1}{n}\mu(A \cap P_n) \geq 0$  since  $\nu_s$  is a positive measure and  $P_n$  is positive for  $\nu_s - \frac{1}{n}\mu$ . Thus we have  $\int_A (f + \frac{1}{n}\mathbf{1}_{P_n}) d\mu \leq \nu(A)$  for every  $A \in \mathcal{S}$ , so  $f + \frac{1}{n}\mathbf{1}_{P_n} \in \mathcal{F}$  for every  $n \in \mathbb{N}_+$ . If  $\mu(P_n) > 0$  then  $\int_S (f + \frac{1}{n}\mathbf{1}_{P_n}) d\mu = \alpha + \frac{1}{n}\mu(P_n) > \alpha$ , which contradicts the definition of  $\alpha$ . Hence we must have  $\mu(P_n) = 0$  for every  $n \in \mathbb{N}_+$ . Now let  $P = \bigcup_{n=1}^{\infty} P_n$ . Then  $\mu(P) = 0$ . If  $\nu_s(P^c) > 0$  then  $\nu_s(P^c) - \frac{1}{n}\mu(P^c) > 0$  for  $n$  sufficiently large. But this is a contradiction since  $P^c \subseteq P_n^c$  which is negative for  $\nu_s - \frac{1}{n}\mu$  for every  $n \in \mathbb{N}_+$ . Thus we must have  $\nu_s(P^c) = 0$ , so  $\mu$  and  $\nu_s$  are singular.

**Part 2.** Suppose that  $\mu$  and  $\nu$  are  $\sigma$ -finite, positive measures. Then there exists a countable partition  $\{S_i : i \in I\}$  of  $S$  where  $S_i \in \mathcal{S}$  for  $i \in I$ , and  $\mu(S_i) < \infty$  and  $\nu(S_i) < \infty$  for  $i \in I$ . Let  $\mu_i(A) = \mu(A \cap S_i)$  and  $\nu_i(A) = \nu(A \cap S_i)$  for  $i \in I$ . Then  $\mu_i$  and  $\nu_i$  are finite, positive measures for  $i \in I$ , and  $\mu = \sum_{i \in I} \mu_i$  and  $\nu = \sum_{i \in I} \nu_i$ . By part 1, for each  $i \in I$ , there exists a measurable function  $f_i : S \rightarrow [0, \infty)$  such that  $\nu_i = \nu_{i,c} + \nu_{i,s}$  where  $\nu_{i,c}(A) = \int_A f_i d\mu$  for  $A \in \mathcal{S}$  and  $\nu_{i,s} \perp \mu$ . Let  $f = \sum_{i \in I} \mathbf{1}_{A_i} f_i$ . Then  $f : S \rightarrow [0, \infty)$  is measurable. Define  $\nu_c(A) = \int_A f d\mu$  and  $\nu_s(A) = \nu(A) - \nu_c(A)$  for  $A \in \mathcal{S}$ . Note that  $\nu_c = \sum_{i \in I} \nu_{i,c}$  and  $\nu_s = \sum_{i \in I} \nu_{i,s}$ . Then  $\nu_c \ll \mu$  and has density function  $f$  and  $\nu_s \perp \mu$ .

**Part 3.** Suppose that  $\nu$  is a  $\sigma$ -finite measure (not necessarily positive). By the Jordan decomposition theorem,  $\nu = \nu_+ - \nu_-$  where  $\nu_+$  and  $\nu_-$  are  $\sigma$ -finite, positive measures, and at least one is finite. By part 2, there exist measurable functions  $f_+ : S \rightarrow [0, \infty)$  and  $f_- : S \rightarrow [0, \infty)$  such that  $\nu_+ = \nu_{+,c} + \nu_{+,s}$  and  $\nu_- = \nu_{-,c} + \nu_{-,s}$  where  $\nu_{+,c}(A) = \int_A f_+ d\mu$ ,  $\nu_{-,c}(A) = \int_A f_- d\mu$  for  $A \in \mathcal{S}$ , and  $\nu_{+,s} \perp \mu$ ,  $\nu_{-,s} \perp \mu$ . Let  $f = f_+ - f_-$ ,  $\nu_c(A) = \int_A f d\mu$ ,  $\nu_s(A) = \nu(A) - \nu_c(A)$  for  $A \in \mathcal{S}$ . Then  $\nu = \nu_c + \nu_s$  and  $\nu_s = \nu_{+,s} - \nu_{-,s} \perp \mu$ .

**Uniqueness.** Suppose that  $\nu = \nu_{c,1} + \nu_{s,1} = \nu_{c,2} + \nu_{s,2}$  where  $\nu_{c,i} \ll \mu$  and  $\nu_{s,i} \perp \mu$  for  $i \in \{1, 2\}$ . Then  $\nu_{c,1} - \nu_{c,2} = \nu_{s,2} - \nu_{s,1}$ . But  $\nu_{c,1} - \nu_{c,2} \ll \mu$  and  $\nu_{s,2} - \nu_{s,1} \perp \mu$  so  $\nu_{c,1} - \nu_{c,2} = \nu_{s,2} - \nu_{s,1} = \mathbf{0}$  by the theorem above.

In particular, a measure  $\nu$  on  $(S, \mathcal{S})$  has a density function with respect to  $\mu$  if and only if  $\nu \ll \mu$ . The density function in this case is also referred to as the *Radon-Nikodym derivative* of  $\nu$  with respect to  $\mu$  and is sometimes written in derivative notation as  $d\nu/d\mu$ . This notation, however, can be a bit misleading because we need to remember that a density function is unique only up to a  $\mu$ -null set. Also, the Radon-Nikodym theorem can fail if the positive measure space  $(S, \mathcal{S}, \mu)$  is not  $\sigma$ -finite. A couple of examples are given below. Next we characterize the Hahn decomposition and the Jordan decomposition of  $\nu$  in terms of the density function.

Suppose that  $\nu$  is a measure on  $(S, \mathcal{S})$  with  $\nu \ll \mu$ , and that  $\nu$  has density function  $f$  with respect to  $\mu$ . Let  $P = \{x \in S : f(x) \geq 0\}$ , and let  $f^+$  and  $f^-$  denote the positive and negative parts of  $f$ .

1. A Hahn decomposition of  $\nu$  is  $(P, P^c)$ .
2. The Jordan decomposition is  $\nu = \nu_+ - \nu_-$  where  $\nu_+(A) = \int_A f^+ d\mu$  and  $\nu_-(A) = \int_A f^- d\mu$ , for  $A \in \mathcal{S}$ .

**Proof**

Of course  $P^c = \{x \in S : f(x) < 0\}$ . The proofs are simple.

1. Suppose that  $A \in \mathcal{S}$ . If  $A \subseteq P$  then  $f(x) \geq 0$  for  $x \in A$  and hence  $\nu(A) = \int_A f d\mu \geq 0$ . If  $A \subseteq P^c$  then  $\nu(A) = \int_A f d\mu \leq 0$ .
2. This follows immediately from (a) and the Jordan decomposition theorem, since  $\nu_+(A) = \nu(A \cap P)$  and  $\nu_-(A) = -\nu(A \cap P^c)$  for  $A \in \mathcal{S}$ . Note that  $f^+ = \mathbf{1}_P f$  and  $f^- = -\mathbf{1}_{P^c} f$ .

The following result is a basic *change of variables theorem* for integrals.

Suppose that  $\nu$  is a positive measure on  $(S, \mathcal{S})$  with  $\nu \ll \mu$  and that  $\nu$  has density function  $f$  with respect to  $\mu$ . If  $g : S \rightarrow \mathbb{R}$  is a measurable function whose integral with respect to  $\nu$  exists, then

$$\int_S g d\nu = \int_S g f d\mu \quad (3.13.6)$$

**Proof**

The proof is a classical *bootstrapping argument*. Suppose first that  $g = \sum_{i \in I} a_i \mathbf{1}_{A_i}$  is a nonnegative simple function. That is,  $I$  is a finite index set,  $a_i \in [0, \infty)$  for  $i \in I$ , and  $\{A_i : i \in I\}$  is a disjoint collection of sets in  $\mathcal{S}$ . Then  $\int_S g d\nu = \sum_{i \in I} a_i \nu(A_i)$ . But  $\nu(A_i) = \int_{A_i} f d\mu = \int_S \mathbf{1}_{A_i} f d\mu$  for each  $i \in I$  so

$$\int_S g d\mu = \sum_{i \in I} a_i \int_S \mathbf{1}_{A_i} f d\mu = \int_S \left( \sum_{i \in I} a_i \mathbf{1}_{A_i} \right) f d\mu = \int_S g f d\mu \quad (3.13.7)$$

Suppose next that  $g : S \rightarrow [0, \infty)$  is measurable. There exists a sequence of nonnegative simple functions  $(g_1, g_2, \dots)$  such that  $g_n$  is increasing in  $n \in \mathbb{N}_+$  on  $S$  and  $g_n \rightarrow g$  as  $n \rightarrow \infty$  on  $S$ . Since  $f$  is nonnegative,  $g_n f$  is increasing in  $n \in \mathbb{N}_+$  on  $S$  and  $g_n f \rightarrow g f$  as  $n \rightarrow \infty$  on  $S$ . By the first step,  $\int_S g_n d\nu = \int_S g_n f d\mu$  for each  $n \in \mathbb{N}_+$ . But by the monotone convergence theorem,  $\int_S g_n d\nu \rightarrow \int_S g d\nu$  and  $\int_S g_n f d\mu \rightarrow \int_S g f d\mu$  as  $n \rightarrow \infty$ . Hence  $\int_S g d\nu = \int_S g f d\mu$ .

Finally, suppose that  $g: S \rightarrow \mathbb{R}$  is a measurable function whose integral with respect to  $\nu$  exists. By the previous step,  $\int_S g^+ d\nu = \int_S g^+ f d\mu$  and  $\int_S g^- d\nu = \int_S g^- f d\mu$ , and at least one of these integrals is finite. Hence by the additive property

$$\int_S g d\nu = \int_S g^+ d\nu - \int_S g^- d\nu = \int_S g^+ f d\mu - \int_S g^- f d\mu = \int_S (g^+ - g^-) f d\mu = \int_S g f d\mu \quad (3.13.8)$$

In *differential notation*, the change of variables theorem has the familiar form  $d\nu = f d\mu$ , and this is really the justification for the derivative notation  $f = d\nu/d\mu$  in the first place. The following result gives the *scalar multiple rule* for density functions.

Suppose that  $\nu$  is a measure on  $(S, \mathcal{S})$  with  $\nu \ll \mu$  and that  $\nu$  has density function  $f$  with respect to  $\mu$ . If  $c \in \mathbb{R}$ , then  $c\nu$  has density function  $cf$  with respect to  $\mu$ .

Proof

If  $A \in \mathcal{S}$  then  $\int_A cf d\mu = c \int_A f d\mu = c\nu(A)$ .

Of course, we already knew that  $\nu \ll \mu$  implies  $c\nu \ll \mu$  for  $c \in \mathbb{R}$ , so the new information is the relation between the density functions. In derivative notation, the scalar multiple rule has the familiar form

$$\frac{d(c\nu)}{d\mu} = c \frac{d\nu}{d\mu} \quad (3.13.9)$$

The following result gives the sum rule for density functions. Recall that two measures are of the *same type* if neither takes the value  $\infty$  or if neither takes the value  $-\infty$ .

Suppose that  $\nu$  and  $\rho$  are measures on  $(S, \mathcal{S})$  of the same type with  $\nu \ll \mu$  and  $\rho \ll \mu$ , and that  $\nu$  and  $\rho$  have density functions  $f$  and  $g$  with respect to  $\mu$ , respectively. Then  $\nu + \rho$  has density function  $f + g$  with respect to  $\mu$ .

Proof

If  $A \in \mathcal{S}$  then

$$\int_A (f + g) d\mu = \int_A f d\mu + \int_A g d\mu = \nu(A) + \rho(A) \quad (3.13.10)$$

The additive property holds because we know that the integrals in the middle of the displayed equation are not of the form  $\infty - \infty$ .

Of course, we already knew that  $\nu \ll \mu$  and  $\rho \ll \mu$  imply  $\nu + \rho \ll \mu$ , so the new information is the relation between the density functions. In derivative notation, the sum rule has the familiar form

$$\frac{d(\nu + \rho)}{d\mu} = \frac{d\nu}{d\mu} + \frac{d\rho}{d\mu} \quad (3.13.11)$$

The following result is the *chain rule* for density functions.

Suppose that  $\nu$  is a positive measure on  $(S, \mathcal{S})$  with  $\nu \ll \mu$  and that  $\nu$  has density function  $f$  with respect to  $\mu$ . Suppose  $\rho$  is a measure on  $(S, \mathcal{S})$  with  $\rho \ll \nu$  and that  $\rho$  has density function  $g$  with respect to  $\nu$ . Then  $\rho$  has density function  $gf$  with respect to  $\mu$ .

Proof

This is a simple consequence of the change of variables theorem above. If  $A \in \mathcal{S}$  then  $\rho(A) = \int_A g d\nu = \int_A gf d\mu$ .

Of course, we already knew that  $\nu \ll \mu$  and  $\rho \ll \nu$  imply  $\rho \ll \mu$ , so once again the new information is the relation between the density functions. In derivative notation, the chain rule has the familiar form

$$\frac{d\rho}{d\mu} = \frac{d\rho}{d\nu} \frac{d\nu}{d\mu} \quad (3.13.12)$$

The following related result is the *inverse rule* for density functions.

Suppose that  $\nu$  is a positive measure on  $(S, \mathcal{S})$  with  $\nu \ll \mu$  and  $\mu \ll \nu$  (so that  $\nu \equiv \mu$ ). If  $\nu$  has density function  $f$  with respect to  $\mu$  then  $\mu$  has density function  $1/f$  with respect to  $\nu$ .

Proof

Let  $f$  be a density function of  $\nu$  with respect to  $\mu$  and let  $Z = \{x \in S : f(x) = 0\}$ . Then  $\nu(Z) = \int_Z f d\mu = 0$  so  $Z$  is a null set of  $\nu$  and hence is also a null set of  $\mu$ . Thus, we can assume that  $f \neq 0$  on  $S$ . Let  $g$  be a density of  $\mu$  with respect to  $\nu$ . Since  $\mu \ll \nu \ll \mu$ , it follows from the chain rule that  $fg$  is a density of  $\mu$  with respect to  $\mu$ . But of course the constant function 1 is also a density of  $\mu$  with respect to itself so we have  $fg = 1$  almost everywhere on  $S$ . Thus  $1/f$  is a density of  $\mu$  with respect to  $\nu$ .

In derivative notation, the inverse rule has the familiar form

$$\frac{d\mu}{d\nu} = \frac{1}{d\nu/d\mu} \quad (3.13.13)$$

## Examples and Special Cases

### Discrete Spaces

Recall that a *discrete measure space*  $(S, \mathcal{S}, \#)$  consists of a countable set  $S$  with the  $\sigma$ -algebra  $\mathcal{S} = \mathcal{P}(S)$  of all subsets of  $S$ , and with counting measure  $\#$ . Of course  $\#$  is a positive measure and is trivially  $\sigma$ -finite since  $S$  is countable. Note also that  $\emptyset$  is the only set that is null for  $\#$ . If  $\nu$  is a measure on  $S$ , then by definition,  $\nu(\emptyset) = 0$ , so  $\nu$  is absolutely continuous relative to  $\mu$ . Thus, by the Radon-Nikodym theorem,  $\nu$  can be written in the form

$$\nu(A) = \sum_{x \in A} f(x), \quad A \subseteq S \quad (3.13.14)$$

for a unique  $f : S \rightarrow \mathbb{R}$ . Of course, this is obvious by a direct argument. If we define  $f(x) = \nu\{x\}$  for  $x \in S$  then the displayed equation follows by the countable additivity of  $\nu$ .

### Spaces Generated by Countable Partitions

We can generalize the last discussion to *spaces generated by countable partitions*. Suppose that  $S$  is a set and that  $\mathcal{A} = \{A_i : i \in I\}$  is a countable partition of  $S$  into nonempty sets. Let  $\mathcal{S} = \sigma(\mathcal{A})$  and recall that every  $A \in \mathcal{S}$  has a unique representation of the form  $A = \bigcup_{j \in J} A_j$  where  $J \subseteq I$ . Suppose now that  $\mu$  is a positive measure on  $\mathcal{S}$  with  $0 < \mu(A_i) < \infty$  for every  $i \in I$ . Then once again, the measure space  $(S, \mathcal{S}, \mu)$  is  $\sigma$ -finite and  $\emptyset$  is the only null set. Hence if  $\nu$  is a measure on  $(S, \mathcal{S})$  then  $\nu$  is absolutely continuous with respect to  $\mu$  and hence has unique density function  $f$  with respect to  $\mu$ :

$$\nu(A) = \int_A f d\mu, \quad A \in \mathcal{S} \quad (3.13.15)$$

Once again, we can construct the density function explicitly.

In the setting above, define  $f : S \rightarrow \mathbb{R}$  by  $f(x) = \nu(A_i)/\mu(A_i)$  for  $x \in A_i$  and  $i \in I$ . Then  $f$  is the density of  $\nu$  with respect to  $\mu$ .

Proof

Suppose that  $A \in \mathcal{S}$  so that  $A = \bigcup_{j \in J} A_j$  for some  $J \subseteq I$ . Then

$$\int_A f d\mu = \sum_{j \in J} \int_{A_j} f d\mu = \sum_{j \in J} \frac{\nu(A_j)}{\mu(A_j)} \mu(A_j) = \sum_{j \in J} \nu(A_j) = \nu(A) \quad (3.13.16)$$

Often positive measure spaces that occur in applications can be decomposed into spaces generated by countable partitions. In the section on Convergence in the chapter on Martingales, we show that more general density functions can be obtained as limits of density functions of the type in the [last theorem](#).

### Probability Spaces

Suppose that  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space and that  $X$  is a random variable taking values in a measurable space  $(S, \mathcal{S})$ . Recall that the *distribution* of  $X$  is the probability measure  $P_X$  on  $(S, \mathcal{S})$  given by

$$P_X(A) = \mathbb{P}(X \in A), \quad A \in \mathcal{S} \quad (3.13.17)$$

If  $\mu$  is a positive measure,  $\sigma$ -finite measure on  $(S, \mathcal{S})$ , then the theory of this section applies, of course. The Radon-Nikodym theorem tells us precisely when (the distribution of)  $X$  has a probability density function with respect to  $\mu$ : we need the distribution to be absolutely continuous with respect to  $\mu$ : if  $\mu(A) = 0$  then  $P_X(A) = \mathbb{P}(X \in A) = 0$  for  $A \in \mathcal{S}$ .

Suppose that  $r : S \rightarrow \mathbb{R}$  is measurable, so that  $r(X)$  is a real-valued random variable. The integral of  $r(X)$  (assuming that it exists) is of fundamental importance, and is known as the *expected value* of  $r(X)$ . We will study expected values in detail in the next chapter, but here we just note different ways to write the integral. By the change of variables theorem in the last section we have

$$\int_{\Omega} r[X(\omega)] d\mathbb{P}(\omega) = \int_S r(x) dP_X(x) \quad (3.13.18)$$

Assuming that  $P_X$ , the distribution of  $X$ , is absolutely continuous with respect to  $\mu$ , with density function  $f$ , we can add to our chain of integrals using Theorem [\(14\)](#):



$$\int_{\Omega} r[X(\omega)] d\mathbb{P}(\omega) = \int_S r(x) dP_X(x) = \int_S r(x) f(x) d\mu(x) \quad (3.13.19)$$

Specializing, suppose that  $(S, \mathcal{S}, \#)$  is a discrete measure space. Thus  $X$  has a discrete distribution and (as noted in the previous subsection), the distribution of  $X$  is absolutely continuous with respect to  $\#$ , with probability density function  $f$  given by  $f(x) = \mathbb{P}(X = x)$  for  $x \in S$ . In this case the integral simplifies:

$$\int_{\Omega} r[X(\omega)] d\mathbb{P}(\omega) = \sum_{x \in S} r(x) f(x) \quad (3.13.20)$$

Recall next that for  $n \in \mathbb{N}_+$ , the  $n$ -dimensional Euclidean measure space is  $(\mathbb{R}^n, \mathcal{R}_n, \lambda_n)$  where  $\mathcal{R}_n$  is the  $\sigma$ -algebra of Lebesgue measurable sets and  $\lambda_n$  is Lebesgue measure. Suppose now that  $S \in \mathcal{R}_n$  and that  $\mathcal{S}$  is the  $\sigma$ -algebra of Lebesgue measurable subsets of  $S$ , and that once again,  $X$  is a random variable with values in  $S$ . By definition,  $X$  has a continuous distribution if  $\mathbb{P}(X = x) = 0$  for  $x \in S$ . But we now know that this is not enough to ensure that the distribution of  $X$  has a density function with respect to  $\lambda_n$ . We need the distribution to be *absolutely* continuous, so that if  $\lambda_n(A) = 0$  then  $\mathbb{P}(X \in A) = 0$  for  $A \in \mathcal{S}$ . Of course  $\lambda_n\{x\} = 0$  for  $x \in S$ , so absolute continuity implies continuity, but not conversely. Continuity of the distribution is a (much) weaker condition than absolute continuity of the distribution. If the distribution of  $X$  is continuous but not absolutely so, then the distribution will not have a density function with respect to  $\lambda_n$ .

For example, suppose that  $\lambda_n(S) = 0$ . Then the distribution of  $X$  and  $\lambda_n$  are mutually singular since  $\mathbb{P}(X \in S) = 1$  and so  $X$  will not have a density function with respect to  $\lambda_n$ . This will always be the case if  $S$  is countable, so that the distribution of  $X$  is discrete. But it is also possible for  $X$  to have a continuous distribution on an uncountable set  $S \in \mathcal{R}_n$  with  $\lambda_n(S) = 0$ . In such a case, the continuous distribution of  $X$  is said to be *degenerate*. There are a couple of natural ways in which this can happen that are illustrated in the following exercises.

Suppose that  $\Theta$  is uniformly distributed on the interval  $[0, 2\pi)$ . Let  $X = \cos \Theta$ ,  $Y = \sin \Theta$ .

1.  $(X, Y)$  has a continuous distribution on the circle  $C = \{(x, y) : x^2 + y^2 = 1\}$ .
2. The distribution of  $(X, Y)$  and  $\lambda_2$  are mutually singular.
3. Find  $\mathbb{P}(Y > X)$ .

Solution

1. If  $(x, y) \in C$  then there exist a unique  $\theta \in [0, 2\pi)$  with  $x = \cos \theta$  and  $y = \sin \theta$ . Hence  $\mathbb{P}[(X, Y) = (x, y)] = \mathbb{P}(\Theta = \theta) = 0$ .
2.  $\mathbb{P}[(X, Y) \in C] = 1$  but  $\lambda_2(C) = 0$ .
3.  $\frac{1}{2}$

The last example is artificial since  $(X, Y)$  has a one-dimensional distribution in a sense, in spite of taking values in  $\mathbb{R}^2$ . And of course  $\Theta$  has a probability density function  $f$  with respect to  $\lambda_1$  given by  $f(\theta) = 1/2\pi$  for  $\theta \in [0, 2\pi)$ .

Suppose that  $X$  is uniformly distributed on the set  $\{0, 1, 2\}$ ,  $Y$  is uniformly distributed on the interval  $[0, 2]$ , and that  $X$  and  $Y$  are independent.

1.  $(X, Y)$  has a continuous distribution on the product set  $S = \{0, 1, 2\} \times [0, 2]$ .
2. The distribution of  $(X, Y)$  and  $\lambda_2$  are mutually singular.
3. Find  $\mathbb{P}(Y > X)$ .

Solution

1. The variables are independent and  $Y$  has a continuous distribution so  $\mathbb{P}[(X, Y) = (x, y)] = \mathbb{P}(X = 2)\mathbb{P}(Y = y) = 0$  for  $(x, y) \in S$ .
2.  $\mathbb{P}[(X, Y) \in S] = 1$  but  $\lambda_2(S) = 0$
3.  $\frac{1}{2}$

The last exercise is artificial since  $X$  has a discrete distribution on  $\{0, 1, 2\}$  (with all subsets measurable and with  $\#$ ), and  $Y$  a continuous distribution on the Euclidean space  $[0, 2]$  (with Lebesgue measurable subsets and with  $\lambda$ ). Both are absolutely continuous;  $X$  has density function  $g$  given by  $g(x) = 1/3$  for  $x \in \{0, 1, 2\}$  and  $Y$  has density function  $h$  given by  $h(y) = 1/2$  for  $y \in [0, 2]$ . So really, the proper measure space on  $S$  is the product measure space formed from these two spaces. Relative to this product space  $(X, Y)$  has a density  $f$  given by  $f(x, y) = 1/6$  for  $(x, y) \in S$ .

It is also possible to have a continuous distribution on  $S \subseteq \mathbb{R}^n$  with  $\lambda_n(S) > 0$ , yet still with no probability density function, a much more interesting situation. We will give a classical construction. Let  $(X_1, X_2, \dots)$  be a sequence of Bernoulli trials with success parameter  $p \in (0, 1)$ . We will indicate the dependence of the probability measure  $\mathbb{P}$  on the parameter  $p$  with a subscript. Thus, we have a sequence of independent indicator variables with

$$\mathbb{P}_p(X_i = 1) = p, \quad \mathbb{P}_p(X_i = 0) = 1 - p \quad (3.13.21)$$

We interpret  $X_i$  as the  $i$ th binary digit (*bit*) of a random variable  $X$  taking values in  $(0, 1)$ . That is,  $X = \sum_{i=1}^{\infty} X_i/2^i$ . Conversely, recall that every number  $x \in (0, 1)$  can be written in binary form as  $x = \sum_{i=1}^{\infty} x_i/2^i$  where  $x_i \in \{0, 1\}$  for each  $i \in \mathbb{N}_+$ . This representation is unique

except when  $x$  is a *binary rational* of the form  $x = k/2^n$  for  $n \in \mathbb{N}_+$  and  $k \in \{1, 3, \dots, 2^n - 1\}$ . In this case, there are two representations, one in which the bits are eventually 0 and one in which the bits are eventually 1. Note, however, that the set of binary rationals is countable. Finally, note that the uniform distribution on  $(0, 1)$  is the same as Lebesgue measure on  $(0, 1)$ .

$X$  has a continuous distribution on  $(0, 1)$  for every value of the parameter  $p \in (0, 1)$ . Moreover,

1. If  $p, q \in (0, 1)$  and  $p \neq q$  then the distribution of  $X$  with parameter  $p$  and the distribution of  $X$  with parameter  $q$  are mutually singular.
2. If  $p = \frac{1}{2}$ ,  $X$  has the uniform distribution on  $(0, 1)$ .
3. If  $p \neq \frac{1}{2}$ , then the distribution of  $X$  is singular with respect to Lebesgue measure on  $(0, 1)$ , and hence has no probability density function in the usual sense.

**Proof**

If  $x \in (0, 1)$  is not a binary rational, then

$$\mathbb{P}_p(X = x) = \mathbb{P}_p(X_i = x_i \text{ for all } i \in \mathbb{N}_+) = \lim_{n \rightarrow \infty} \mathbb{P}_p(X_i = x_i \text{ for } i = 1, 2, \dots, n) = \lim_{n \rightarrow \infty} p^y (1-p)^{n-y} \quad (3.13.22)$$

where  $y = \sum_{i=1}^n x_i$ . Let  $q = \max\{p, 1-p\}$ . Then  $p^y (1-p)^{n-y} \leq q^n \rightarrow 0$  as  $n \rightarrow \infty$ . Hence,  $\mathbb{P}_p(X = x) = 0$ . If  $x \in (0, 1)$  is a binary rational, then there are two bit strings that represent  $x$ , say  $(x_1, x_2, \dots)$  (with bits eventually 0) and  $(y_1, y_2, \dots)$  (with bits eventually 1). Hence  $\mathbb{P}_p(X = x) = \mathbb{P}_p(X_i = x_i \text{ for all } i \in \mathbb{N}_+) + \mathbb{P}_p(X_i = y_i \text{ for all } i \in \mathbb{N}_+)$ . But both of these probabilities are 0 by the same argument as before.

Next, we define the set of numbers for which the limiting relative frequency of 1's is  $p$ . Let  $C_p = \{x \in (0, 1) : \frac{1}{n} \sum_{i=1}^n x_i \rightarrow p \text{ as } n \rightarrow \infty\}$ . Note that since limits are unique,  $C_p \cap C_q = \emptyset$  for  $p \neq q$ . Next, by the strong law of large numbers,  $\mathbb{P}_p(X \in C_p) = 1$ . Although we have not yet studied the law of large numbers, the basic idea is simple: in a sequence of Bernoulli trials with success probability  $p$ , the long-term relative frequency of successes is  $p$ . Thus the distributions of  $X$ , as  $p$  varies from 0 to 1, are mutually singular; that is, as  $p$  varies,  $X$  takes values with probability 1 in mutually disjoint sets.

Let  $F$  denote the distribution function of  $X$ , so that  $F(x) = \mathbb{P}_p(X \leq x) = \mathbb{P}_p(X < x)$  for  $x \in (0, 1)$ . If  $x \in (0, 1)$  is not a binary rational, then  $X < x$  if and only if there exists  $n \in \mathbb{N}_+$  such that  $X_i = x_i$  for  $i \in \{1, 2, \dots, n-1\}$  and  $X_n = 0$  while  $x_n = 1$ . Hence  $\mathbb{P}_{1/2}(X < x) = \sum_{n=1}^{\infty} \frac{x_n}{2^n} = x$ . Since the distribution function of a continuous distribution is continuous, it follows that  $F(x) = x$  for all  $x \in [0, 1]$ . This means that  $X$  has the uniform distribution on  $(0, 1)$ . If  $p \neq \frac{1}{2}$ , the distribution of  $X$  and the uniform distribution are mutually singular, so in particular,  $X$  does not have a probability density function with respect to Lebesgue measure.

For an application of some of the ideas in this example, see Bold Play in the game of Red and Black.

### Counterexamples

The essential uniqueness of density functions can fail if the underlying positive measure  $\mu$  is not  $\sigma$ -finite. Here is a trivial counterexample:

Suppose that  $S$  is a nonempty set and that  $\mathcal{S} = \{S, \emptyset\}$  is the trivial  $\sigma$ -algebra. Define the positive measure  $\mu$  on  $(S, \mathcal{S})$  by  $\mu(\emptyset) = 0$ ,  $\mu(S) = \infty$ . Let  $\nu_c$  denote the measure on  $(S, \mathcal{S})$  with constant density function  $c \in \mathbb{R}$  with respect to  $\mu$ .

1.  $(S, \mathcal{S}, \mu)$  is not  $\sigma$ -finite.
2.  $\nu_c = \mu$  for every  $c \in (0, \infty)$ .

The Radon-Nikodym theorem can fail if the measure  $\mu$  is not  $\sigma$ -finite, even if  $\nu$  is finite. Here are a couple of standard counterexamples:

Suppose that  $S$  is an uncountable set and  $\mathcal{S}$  is the  $\sigma$ -algebra of *countable and co-countable sets*:

$$\mathcal{S} = \{A \subseteq S : A \text{ is countable or } A^c \text{ is countable}\} \quad (3.13.23)$$

As usual, let  $\#$  denote counting measure on  $\mathcal{S}$ , and define  $\nu$  on  $\mathcal{S}$  by  $\nu(A) = 0$  if  $A$  is countable and  $\nu(A) = 1$  if  $A^c$  is countable. Then

1.  $(S, \mathcal{S}, \#)$  is not  $\sigma$ -finite.
2.  $\nu$  is a finite, positive measure on  $(S, \mathcal{S})$ .
3.  $\nu$  is absolutely continuous with respect to  $\#$ .
4.  $\nu$  does not have a density function with respect to  $\#$ .

**Proof**

1. Recall that a countable union of countable sets is countable, and so  $S$  cannot be written as such a union.
2. Note that  $\nu(\emptyset) = 0$ . Suppose that  $\{A_i : i \in I\}$  is a countable, disjoint collection of sets in  $\mathcal{S}$ . If  $A_i$  is countable for every  $i \in I$  then  $\bigcup_{i \in I} A_i$  is countable. Hence  $\nu(\bigcup_{i \in I} A_i) = 0$  and  $\nu(A_i) = 0$  for every  $i \in I$ . Next suppose that  $A_j^c$  and  $A_k^c$  are countable for distinct  $j, k \in I$ . Since  $A_j \cap A_k = \emptyset$ , we have  $A_j^c \cup A_k^c = S$ . But then  $S$  would be countable, which is a contradiction. Hence it is only possible for to have  $A_j^c$  countable for a single  $j \in I$ . In this case,  $\nu(A_j) = 1$  and  $\nu(A_i) = 0$  for  $i \neq j$ . But also  $(\bigcup_{i \in I} A_i)^c = \bigcap_{i \in I} A_i^c$



is countable, so  $\nu(\bigcup_{i \in I} A_i) = 1$ . Hence in all cases,  $\nu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \nu(A_i)$  so  $\nu$  is a measure on  $(S, \mathcal{S})$ . It is clearly positive and finite.

3. Recall that any measure is absolutely continuous with respect to counting measure, since  $\#(A) = 0$  if and only if  $A = \emptyset$ .
4. Suppose that  $\nu$  has density function  $f$  with respect to  $\#$ . Then  $0 = \nu\{x\} = \int_{\{x\}} f d\# = f(x)$  for every  $x \in S$ . But then  $\nu(S) = \int_S f d\# = 0$ , which is a contradiction.

Let  $\mathcal{R}$  denote the standard Borel  $\sigma$ -algebra on  $\mathbb{R}$ . Let  $\#$  and  $\lambda$  denote counting measure and Lebesgue measure on  $(\mathbb{R}, \mathcal{R})$ , respectively. Then

1.  $(\mathbb{R}, \mathcal{R}, \#)$  is not  $\sigma$ -finite.
2.  $\lambda$  is absolutely continuous with respect to  $\#$ .
3.  $\lambda$  does not have a density function with respect to  $\#$ .

Proof

1.  $\mathbb{R}$  is uncountable and hence cannot be written as a countable union of finite sets.
2. Since  $\emptyset$  is the only null set of  $\#$ ,  $\lambda \ll \#$ .
3. Suppose that  $\lambda$  has density function  $f$  with respect to  $\#$ . Then

$$0 = \lambda\{x\} = \int_{\{x\}} f d\# = f(x), \quad x \in \mathbb{R} \quad (3.13.24)$$

But then also  $\lambda(\mathbb{R}) = \int_{\mathbb{R}} f d\# = 0$ , a contradiction.

This page titled [3.13: Absolute Continuity and Density Functions](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.