We apply the martingale convergence theorem to prove the Radon-Nikodym theorem, which states that if $\mu$ and $\nu$ are $\sigma$-finite measures on a measurable space $(\Omega, \mathcal{F})$ and $\nu$ is absolutely continuous with respect to $\mu$ then there exists a non-negative and measurable $f : \Omega \to \mathbb{R}$ such that $\nu(A) = \int_A f \, d\mu$ for all measurable sets $A$.

As any $\sigma$-finite measure is equivalent to a probability measure, it is enough to prove the result in the case where $\mu$ and $\nu$ are probability measures. Furthermore, by the Jordan decomposition, the result generalizes to the case where $\nu$ is a signed measure. So, we just need to prove the following.

**Theorem (Radon-Nikodym).** Let $P$ and $Q$ be probability measures on the measurable space $(\Omega, \mathcal{F})$, such that $Q$ is absolutely continuous with respect to $P$. Then, there exists a non-negative random variable $X$ such that $E_P[X] = 1$ and $Q(A) = E_P[1_A X]$ for every $A \in \mathcal{F}$.

Here, $X$ is called the Radon-Nikodym derivative of $Q$ with respect to $P$.

More generally, for any sub-$\sigma$-algebra $\mathcal{G}$ of $\mathcal{F}$ we can restrict the measures $P$ and $Q$ to $\mathcal{G}$ and ask if the Radon-Nikodym derivative of $Q|_G$ with respect to $P|_G$ exists. If it does we shall denote it by $X_G$, which by definition is a non-negative $\mathcal{G}$-measurable random variable satisfying $Q(A) = E_P[1_A X_G]$ for all $A \in \mathcal{G}$.

We note that if $X_G$ does exist, then it is uniquely defined ($P$-almost everywhere). Suppose that $\hat{X}_G$ also satisfied the required properties, then

$$E_P[\max(X_G - \hat{X}_G, 0)] = E_P[X_G 1_{\{X_G > \hat{X}_G\}}] - E_P[\hat{X}_G 1_{\{X_G > \hat{X}_G\}}] = 0$$

so $X_G \leq \hat{X}_G$ almost surely. Similarly, $\hat{X}_G \leq X_G$ and therefore $\hat{X}_G = X_G$ (almost surely).

First, the easy case. For a finite $\sigma$-algebra, the Radon-Nikodym derivative can be written out explicitly.

---

*({MartingaleProofOfTheRadonNikodymTheorem} created: 2013-03-2 by: {gel} version: ⟨4129⟩ Privacy setting: ⟨1⟩ {Proof} ⟨60G12⟩ ⟨28A15⟩

†This text is available under the Creative Commons Attribution/Share-Alike License 3.0. You can reuse this document or portions thereof only if you do so under terms that are compatible with the CC-BY-SA license.
Lemma 1. If \( \mathcal{G} \) is a finite sub-\( \sigma \)-algebra of \( \mathcal{F} \) then the Radon-Nikodym derivative \( X_\mathcal{G} \) exists.

Proof. Let \( A_1, A_2, \ldots, A_n \) be the minimal non-empty elements of \( \mathcal{G} \). These are pairwise disjoint subsets of \( \Omega \) such that every set in \( \mathcal{G} \) is a union of a subcollection of the \( A_k \). Set

\[
X_\mathcal{G} = \sum_{k=1}^{n} \frac{Q(A_k)}{P(A_k)} 1_{A_k}
\]

Note that whenever \( P(A_k) = 0 \) then \( Q(A_k) = 0 \), and we adopt the convention that \( \frac{0}{0} = 0 \). Clearly, \( X_\mathcal{G} \) is \( \mathcal{G} \)-measurable, and

\[
E_P[1_{A_k} X_\mathcal{G}] = \frac{Q(A_k)}{P(A_k)} E_P[1_{A_k}] + \sum_{j \neq k} \frac{Q(A_j)}{P(A_j)} E_P[1_{A_k \cap A_j}]
\]

\[
= Q(A_k).
\]

Here, we have used \( E_P[1_{A_k}] = P(A_k) \) and \( 1_{A_k \cap A_j} = 0 \). By linearity, this equality remains true if both sides are replaced by any union of the \( A_k \), and therefore \( X_\mathcal{G} \) is the required Radon-Nikodym derivative.

Next, martingale convergence is used to prove the existence of the Radon-Nikodym derivative in the case where the \( \sigma \)-algebra \( \mathcal{G} \) is separable. By separable, we mean that there is a countable sequence of sets \( A_1, A_2, \ldots \) generating \( \mathcal{G} \). Note that if we let \( \mathcal{G}_n \) be the \( \sigma \)-algebra generated by \( A_1, A_2, \ldots, A_n \), then \( \mathcal{G}_n \) is an increasing sequence of finite sub-\( \sigma \)-algebras such that \( \bigcup_n \mathcal{G}_n \) generates \( \mathcal{G} \). The following result is general enough to apply in many useful cases, such as with the Boral \( \sigma \)-algebra on \( \mathbb{R}^n \).

Lemma 2. Let \( \mathcal{G} \) be a separable sub-\( \sigma \)-algebra of \( \mathcal{F} \). Then, the Radon-Nikodym derivative \( X_\mathcal{G} \) exists. If furthermore, \( \mathcal{G}_n \) is an increasing sequence of finite \( \sigma \)-algebras satisfying \( \mathcal{G} = \sigma(\bigcup_n \mathcal{G}_n) \) then \( E_P[|X_\mathcal{G} - X_{\mathcal{G}_n}|] \to 0 \) as \( n \to \infty \).

Proof. Let us set \( X_n \equiv X_{\mathcal{G}_n} \). If \( m < n \) then the conditional expectation \( E_P[X_n | \mathcal{G}_m] \) is \( \mathcal{G}_m \)-measurable, and for every \( A \in \mathcal{G}_m \),

\[
E_P[1_A E_P[X_n | \mathcal{G}_m]] = E_P[1_A X_n] = Q(A).
\]

This equality just uses the definition of the conditional expectation and then the definition of \( X_n \) as the Radon-Nikodym derivative restricted to \( \mathcal{G}_n \). So, \( E_P[X_n | \mathcal{G}_m] \) is the Radon-Nikodym derivative restricted to \( \mathcal{G}_m \), and equals \( X_m \) (almost-surely).

Therefore, \( X_n \) is a martingale and the martingale convergence theorem implies that the limit

\[
X_\mathcal{G} = \lim_{n \to \infty} X_n
\]

exists almost surely. We now show that the sequence \( X_n \) is uniformly integrable. Choose any \( \epsilon > 0 \). As \( Q \) is absolutely continuous with respect to \( P \), there exists
a $\delta > 0$ such that $Q(A) < \epsilon$ whenever $P(A) < \delta$. Using

$$P(X_n > K) = E_P[1_{\{X_n > K\}}] \leq E_P\left(\frac{X_n}{K}\right) = \frac{1}{K}$$

we see that $P(X_n > K) < \delta$ whenever $K > \delta^{-1}$ and, therefore, $Q(X_n > K) < \epsilon$. So

$$E_P[X_n 1_{\{X_n > K\}}] = Q(X_n > K) < \epsilon$$

for every $n$, showing that $X_n$ is a uniformly integrable sequence with respect to $P$. Therefore, convergence in ($\pi$-system, $\pi$-system, $\pi$-system) is a Dynkin system containing the $\pi$-system $\bigcup_n \mathcal{G}_n$ so, by Dynkin’s lemma, is satisfied for every $A \in \sigma(\bigcup_n \mathcal{G}_n) = \mathcal{G}$ and, by definition, $X_\mathcal{G}$ is the Radon-Nikodym derivative restricted to $\mathcal{G}$.

Finally, by approximating by finite $\sigma$-algebras we can prove the Radon-Nikodym theorem for arbitrary inseparable $\sigma$-algebras $\mathcal{F}$.

**Proof of the Radon-Nikodym theorem:**

First, we use contradiction to show that for any $\epsilon > 0$ there exists a finite $\sigma$-algebra $\mathcal{G} \subseteq \mathcal{F}$ satisfying $E_P[|X_\mathcal{G} - X_H|] < \epsilon$ for every finite $\sigma$-algebra $H$ with $\mathcal{G} \subseteq H \subseteq F$. If this were not the case, then by induction we could find an increasing sequence of finite $\sigma$-algebras of $\mathcal{F}$ satisfying $E_P[|X_{\mathcal{G}_n} - X_{\mathcal{G}_{n+1}}|] \geq \epsilon$. However, letting $\mathcal{G} = \sigma(\bigcup_n \mathcal{G}_n)$, Lemma ?? shows that $X_\mathcal{G}$ exists and

$$\epsilon \leq \lim_{n \to \infty} E_P[|X_{\mathcal{G}_n} - X_{\mathcal{G}_{n+1}}|] \leq \lim_{n \to \infty} E_P[|X_{\mathcal{G}_n} - X_\mathcal{G}|] + \lim_{n \to \infty} E_P[|X_{\mathcal{G}_{n+1}} - X_\mathcal{G}|] = 0$$

—a contradiction.

So, there exists a sequence of finite $\sigma$-algebras $\mathcal{G}_n$ of $\mathcal{F}$ such that $E_P[|X_{\mathcal{G}_n} - X_\mathcal{H}|] < 2^{-n}$ for every finite $\sigma$-algebra $\mathcal{H}$ of $\mathcal{F}$ containing $\mathcal{G}_n$. Let $\mathcal{G}$ be the (separable) $\sigma$-algebra generated by $\bigcup_n \mathcal{G}_n$, and set $\mathcal{G}_n = \sigma(\bigcup_{k=1}^n \mathcal{G}_k)$. By Lemma ??, the Radon-Nikodym derivative restricted to $\mathcal{G}$, $X_\mathcal{G}$, exists, and we show that it is the required derivative of $Q$ with respect to $P$.

Choose any set $A \in \mathcal{F}$ and let $\mathcal{H}_n$ be the (finite) $\sigma$-algebra generated by $\mathcal{G}_n \cup \{A\}$. Then, $X_{\mathcal{H}_n}$ exists and satisfies $E_P[X_{\mathcal{H}_n} 1_A] = Q(A)$ and,

$$|E_P[X_\mathcal{G} 1_A] - Q(A)| = \lim_{n \to \infty} |E_P[X_{\mathcal{G}_n} 1_A] - Q(A)|$$

$$= \lim_{n \to \infty} |E_P[X_{\mathcal{G}_n} 1_A] - E_P[X_{\mathcal{H}_n} 1_A]|$$

$$\leq \lim_{n \to \infty} E_P[|X_{\mathcal{G}_n} - X_{\mathcal{H}_n}|] + \lim_{n \to \infty} E_P[|X_{\mathcal{H}_n} - X_{\mathcal{G}_n}|]$$

$$\leq \lim_{n \to \infty} (2^{-n} + 2^{-n}) = 0.$$

So, $E_P[X_\mathcal{G} 1_A] = Q(A)$ as required.
References