Definition: density

A r.v. X and the corresponding distribution P_X has a **density** f wrt Lebesgue measure if

(i) $f : \mathbb{R} \to [0, \infty]$ is measurable $\mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R}^+ \cup \{+\infty\})$.

(ii) For all $A \in \mathcal{B}(\mathbb{R})$,

$$P_X(A) = P(X \in A) = \int_A f(x)dx.$$

Note.

1) can define wrt other measures.
2) we have the familiar

$$\int_{\mathbb{R}} f dx = P_X(\mathbb{R}) = p(X \in \mathbb{R}) = P(\Omega) = 1.$$

3) A r.v. X may have multiple densities wrt Lebesgue measure. They can only differ on sets with Lebesgue measure zero.

Claim. Define $F(x) = P(X \leq x) = P(X \in (-\infty, x]) = \int_{-\infty}^{x} f(t)dt$. If F is a cdf (right-continuous, etc), then this equation implies that F is continuous.

Consider the random vector $X = (X_1, \ldots, X_k)$.

1) This induces a probability measure $P_X(A) = P((X_1, \ldots, X_k) \in A)\ \forall\ A \in \mathcal{B}(\mathbb{R}^k)$.
2) has a distribution function

$$F_X(x_1, \ldots, x_k) \coloneqq P(X_1 \leq x_1, \ldots, X_k \leq x_k).$$

3) may or may not have a density f satisfying $P_X(A) = \int_A f(t)dt$.

$$F_X(x_1, \ldots, x_k) = \int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_1} f(t_1, \ldots, t_k)dt_1 \ldots dt_k$$

Property.

1) non-decreasing on each variable (intersections).
2) right-continuous in each variable. That is,

$$\lim_{x_i \searrow a} F_X(x_1, \ldots, x_k) = F(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_k)..$$

When $k = 2$, this is often referred to as "continuity from the right and from above".

3) $\lim_{x_i \to -\infty} F_X(x_1, \ldots, x_k) = 0.$

4) $\lim_{x_i \to \infty} F_X(x_1, \ldots, x_k) = F_X(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$. This is the marginal CDF. Taking all $x_i \to \infty$ gives us 1.

5) F_X can have an uncountable number of discontinuities.

   Example. In $\mathbb{R}^2$, Extend Lebesgue measure on $[0, 1]$ to $\mathbb{R}^2$ as follows:
   Let $A \in \mathcal{B}(\mathbb{R}^2)$, $X = (X_1, X_2)$. Define

$$P_X(A) = \lambda(A \cap \{(x, y) : 0 \leq x \leq 1, y = 0\}).$$

   It's the length of the overlap of A with $[0, 1]$. Note that P_X is a probability measure on $\mathbb{R}^2$ because it is between 0 and 1, $P_X(\emptyset) = 0$, and $P_X(\mathbb{R}^2) = \lambda([0, 1]) = 1$. And given $A_1, \ldots$ disjoint, then

$$\begin{aligned} P_X\left(\bigcup_{n=1}^{\infty} A_n\right) &= \lambda\left(\left(\bigcup_{n=1}^{\infty} A_n\right) \cap [0, 1]\right) \\ &= \lambda\left(\bigcup_{n=1}^{\infty} (A_n \cap [0, 1])\right) \\ &= \sum_{n=1}^{\infty} \lambda(A_n \cap [0, 1]) \text{ disjoint} \\ &= \sum_{n=1}^{\infty} P_X(A_n) \end{aligned}$$

   Now define $F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$. Then

$$\begin{aligned} F(1, 0) &= P(X_1 \leq 1, X_2 \leq 0) \\ &= P_X((-\infty, 1] \times (-\infty, 0]) \\ &= \lambda([0, 1]) = 1 \end{aligned}$$

   Notice

$$F(1, 0^-) = P(X_1 \leq 1, X_2 < 0) = 0.$$

   which is a jumped discontinuity. In fact, $F(a, 0)$ and $F(a, 0^-)$ for any $a \in [0, 1]$ will give us a jump discontinuity. And this is uncountably many of a.

6) In general, since cdfs are right-continuous, F is continuous at $x \Rightarrow F$ is right continuous at $x \Rightarrow$ the boundary of the half rectangle $\{y : y_i \leq x_i, i = 1, 2, \ldots, k\}$ has probability measure zero.

7) OTOH, while F may have an uncountably infinite number of discontinuities, the points of continuities are a dense set in $\mathbb{R}^k$.

Proof

Sketch: Take any $x \in \mathbb{R}^k$ and consider the half rectangles $\{y_i : y_i \leq x_i + h, i = 1, \ldots, k\}$ for various h. The boundaries of these half rectangles are disjoint. Recall from Section 10 if a measure μ on a σ-field $\mathcal{F}$ is σ-finite, then $\mathcal{F}$ cannot contain an uncountable disjoint collection of sets with positive μ measure. P_X is finite $\Rightarrow$ P_X is σ-finite, so only countably many of the $x + h$ rectangle boundaries can have positive P_X measure. That is, only countably many of the $x + h$'s can be points of discontinuities for F. Therefore, we can always select points of continuity of F that approach x since h is arbitrary. $\square$