Section 1.7. Continuous Random Variables

Note. We now formally define more of the ideas introduced in Section 1.5.

Definition 1.7.1. A random variable X is a *continuous random variable* if its cumulative distribution function F_X is a continuous function for all $x \in \mathbb{R}$.

Note. By Theorem 1.5.3, for random variable X we have

$$P(X = x) = F_X(x) - F_X(x^-) = F_X(x) - \lim_{z \to x^-} F_X(z) \text{ for all } x \in \mathbb{R}.$$

So if the cumulative distribution function F_X is continuous that $\lim_{z\to x^-} F_X(z) = F_X(x)$ and $P(X = x) = 0$. That is, for continuous random variable X we have $P(X = x) = 0$ for all $x \in \mathbb{R}$.

Definition. If for continuous random variable X we have that the cumulative distribution function F_X satisfies $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$ for some function f_X, then f_X is the *probability density function* (*pdf*) of X. In this case, the *support* of X is $\mathcal{S} = \{x \in \mathbb{R} \mid f_X(x) > 0\}$.

Note 1.7.A. If the probability density function f_X of continuous random variable X is itself continuous, then by the Fundamental Theorem of Calculus (see my online notes for Calculus 1 on 5.4. The Fundamental Theorem of Calculus) we have

$$\frac{d}{dx}[F_X(x)] = \frac{d}{dx}\left[\int_{-\infty}^{x} f_X(t)\,dt\right] = f_X(x).$$

Note. The text book mentions absolute continuity on page 49. If we use Lebesgue integration instead of Riemann integration then we can get a lot of use out of absolute continuity and we can even generalize the previous note. We now explore some of this, which is covered in Real Analysis 1 (MATH 5210). The following two definitions and one theorem are based on my online notes for 6.4. Absolutely Continuous Functions and 6.5. Integrating Derivatives: Differentiating Indefinite Integrals.

Definition. A real-valued function f on a closed, bounded interval $[a, b]$ is *absolutely continuous* on $[a, b]$ if for each $\varepsilon > 0$ there is $\delta > 0$ such that for every finite disjoint collection $\{(a_k, b_k)\}_{k=1}^n$ of open intervals in (a, b),

$$\text{if } \sum_{k=1}^n (b_k - a_k) < \delta \text{ then } \sum_{k=1}^n |f(b_k) - f(a_k)| < \varepsilon.$$

Definition. A function f on closed, bounded interval $[a, b]$ is the *indefinite integral* of g over $[a, b]$ if g is Lebesgue integrable over $[a, b]$ and

$$f(x) = f(a) + \int_a^x g \text{ for all } x \in [a, b].$$

Theorem 6.11. A function f on a closed, bounded interval $[a, b]$ is absolutely continuous on $[a, b]$ if and only if it is an indefinite integral over $[a, b]$.

Note. We can conclude that if the cumulative distribution function F_X of continuous random variable X satisfies the ε/δ absolute continuity definition given above, then $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$ for some probability density function f_X. Alternatively, we can follow the approach commonly taken in a Measure Theory Based Probability class (which ETSU does not have) where a random variable X is defined as absolutely continuous if there is some nonnegative function f_X (a "Borel measurable" function) defined on $\mathbb{R}$ such that

$$P(X \leq x) = F_X(x) = \int_{-\infty}^{x} f(t)\,dt \text{ for all } x \in \mathbb{R}$$

where the integral here is the Lebesgue integral (see my online notes on 4.6. Random Variables). In <u>this</u> class, we overlook these subtleties and most of our examples will involve Riemann integrals.

Note. If X is a continuous random variable with probability density function f_X then we can calculate the probability that X lies in an interval as follows:

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$$
$$= F_X(b) - F_X(a) = \int_{a}^{b} f_X(t)\,dt.$$

We can similarly find the probability that X lies in a union of intervals.

Example 1.7.1. Suppose a point is selected at random in the interior of a circle of radius 1 in such a way that the probability that the point lies in an open disk D of area A is $P(D) = \text{area}(A)/\pi$. Let X be the distance of the selected point from the origin. The sample space is $\mathcal{C} = \{(w, y) \mid x^2 + y^2 < 1\}$. For $0 < x < 1$, the event $\{X \leq x\}$ corresponds to the point lying in a circle of radius x centered at the origin. So $P(X \leq x) = \pi x^2/\pi = x^2$ and hence the cumulative distribution function of X is

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ x^2 & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 1. \end{cases}$$

So the probability density function satisfies

$$f_X(x) = \frac{d}{dz}[F_X(x)] = \begin{cases} 2x & \text{if } 0 \leq x < 1 \\ 0 & \text{if } x > 1. \end{cases}$$

Notice that $\frac{d}{dx}[F_X(x)]$ is undefined at $x = 1$, but for a continuous random variable, F_X is unaffected by the value of f_X at a finite number of points so we take $f_X(1) = 0$ and then

$$f_X(x) = \begin{cases} 2x & \text{if } 0 \leq x < 1 \\ 0 & \text{elsewhere.} \end{cases}$$

Definition 1.7.2. Let $0 < p < 1$. A *quantile of order* p of the cumulative distribution function F_X of random variable X is a value ξ_p such that $P(X < \xi_p) \leq p$ and $P(X \leq \xi_p) \geq p$. If is also called the $(100p)$*th percentile* of X.

Note. For X a continuous random variable we have $P(X = \xi_p) = 0$ so that $P(X < \xi_p) = P(X \leq \xi_p)$ and for ξ_p a quantile of p then we have both $P(X \leq \xi_p) \leq p$ and $P(\xi_p \geq p$ so that $P(X \leq \xi_p) = p$. This behavior need not be the case for a discrete random variable.

Note. The definition of quantile of order p refers to "a" value ξ_p. This value may not be unique, as the follow example shows.

Example. Consider the probability density function

$$f_X(x) = \begin{cases} 2(1-x)(x-2) & \text{for } 1 < x < 2 \\ 3(3-x)(x-4) & \text{for } 3 < x < 4 \\ 0 & \text{elsewhere.} \end{cases}$$

For $\xi \in [2,3]$ we have

$$\begin{aligned} P(X < \xi) &= P(X \le \xi) \text{ since } P(X = \xi) = 0 \text{ because} \\ &\quad \text{we have a continuous random variable} \\ &= F_X(\xi) = \int_{-\infty}^{\xi} f_X(x)\,dx \\ &= \int_{-\infty}^{1} f_X(x)\,dx + \int_{1}^{2} f_X(x)\,dx + \int_{2}^{\xi} f_X(x)\,dx \\ &= \int_{-\infty}^{1} 0\,dx + \int_{1}^{2} (-3x^2 + 9x - 6)\,dx + \int_{x}^{\xi} 0\,dx \\ &= \left(-x^3 + \frac{9}{2}x^2 - 6x\right)\Big|_1^2 = (-8 + 18 - 12) - (-1 + 9/2 - 6) = 1/2. \end{aligned}$$

So $P(X < \xi) \le 1/2 = p$ and $P(\xi \le \xi_p) \ge 1/2 = p$. Therefore a quantile of order $p = 1/2$ is ξ for any $\xi \in [2,3]$.

Definition. A *median* of a random variable X is a quantile of order $p = 1/2$, $\xi_{1/2}$. A *first quartile* is a quantile of order $p = 1/4$, $\xi_{1/4}$, and a third quartile is a quantile of order $p = 3/4$, $\xi_{1/4}$. A difference $\xi_{3/4} - \xi_{1/4}$ is an *interquartile range* of X.

Note. A median is often used as a measure of the center of the distribution of X and an interquartile range is used as a measure of the spread of the distribution of X.

Note. We can transform a continuous random variable, just as we transformed a discrete random variable. We illustrate this with some examples.

Example 1.7.4. Consider again Example 1.7.1 where a point is selected in an open unit disk and the cumulative distribution function of random variable X, the distance of the point from the center of the disk, is

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ x^2 & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 1. \end{cases}$$

Let Y be the distance square of the point from the center of the circle so that $Y = X^2$. Since the support of X is $\mathcal{S}_X = (0,1)$ then the support of Y is $\mathcal{S}_Y = (0,1)$. The cumulative distribution function of Y satisfies

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(X \leq \sqrt{y}) = F_X(\sqrt{y}) = \begin{cases} 0 & \text{for } y < 0 \\ y & \text{for } 0 \leq y \leq 1 \\ 1 & \text{for } y > 1. \end{cases}$$

So the probability density function of Y is

$$f_X(x) = \begin{cases} 1 & \text{for } 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Example 1.7.5. Let $f_X(x) = \begin{cases} 1/2 & \text{for } -1 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$ be the probability density function of random variable X. Define the random variable $Y = X^2$. If $0 \le y \le 1$ then we have

$$\begin{aligned} P(Y \le y) &= P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) \\ &= P(X \le \sqrt{y}) - P(X < -\sqrt{y}) \\ &= P(X \le \sqrt{y}) - P(X \le -\sqrt{y}) \text{ since } P(X = \sqrt{y}) = 0 \\ &= F_X(\sqrt{y} - F_X(-\sqrt{y}) = \int_{x=-\infty}^{x=\sqrt{y}} \frac{1}{2}\, dx - \int_{-\infty}^{x=-\sqrt{y}} \frac{1}{2}\, dx \\ &= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\, dx = \sqrt{y}. \end{aligned}$$

If $y < 0$ then $P(Y \le y) = 0$ and if $y > 1$ then $P(Y \le y) = 1$. That is,

$$P(Y \le y) = F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ \sqrt{y} & \text{for } 0 \le y \le 1 \\ 1 & \text{for } 1 < y. \end{cases}$$

So the probability density function of Y is

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{y}} & \text{for } 0 < y \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Note. In the previous two examples, we have the transformation $g(x) = x^2$ so that $Y = g(X) = X^2$. In Example 1.7.5 g is not one-to-one so that g^{-1} does not exist. However, in Example 1.7.4 g is one-to-one on the support of X so that we could find the cumulative distribution function of Y using g^{-1}, as described in the next theorem.

Theorem 1.7.1. The Cumulative Distribution Function Technique. Let X be a continuous random variable with probability density function f_X and support $\mathcal{S}_X$. Let $Y = g(X)$ where g is a one-to-one differentiable function on the $x = g^{-1}(y)$ and let $dx/dy = \frac{d}{dy}[g^{-1}(y)]$. Then the probability density function of Y is given by $f_Y(y) = f_X(g^{-1}(y))|dx/dy|$ for $y \in \mathcal{S}_Y$ where the support of Y is the set $\mathcal{S}_Y = \{y = g(x) \mid x \in \mathcal{S}_X\}$.

Note. The test refers to the quantity $dx/dy = d[g^{-1}(y)]/dx$ in Theorem 1.7.1 as the "Jacobian." However, this term is more commonly used in the setting of functions of several variables; see my online notes for Calculus 3 (MATH 2110) on 15.8. Substitution in Multiple Integrals.

Note. Theorem 1.7.1 can be expressed as an algorithm for finding the probability density function of Y where $Y = g(X)$ for one-to-one differentiable g as follows:

1. Find the support of Y.

2. Solve $y = g(x)$ for $x = g^{-1}(y)$.

3. Calculate $dx/dy = d]g^{-1}(y)]/dy$.

4. The probability density function of Y is $f_Y(y) = f_X(g^{-1}(y)) \left|\frac{dx}{dy}\right|$.

Example. Exercise 1.7.24.

Note. For random variable X, we can have a cumulative distribution function that is a mixture of discrete and continuous distributions. By Theorem 1.5.1(d), every cumulative distribution function is right continuous and by Theorem 1.5.3 $P(X = x) = F_X(x) - \lim_{z\to x^-} F_X(z)$ So we can combine these two behaviors to get a mixed cumulative density function.

Example 1.7.7. Consider the cumulative density function

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ (x+1)/2 & \text{for } 0 \le x < 1 \\ 1 & \text{for } 1 \le x. \end{cases}$$

Then $P(X = 0) = F_X(0) - \lim_{z\to z^-} F_X(z) = 1/2 - 0 = 1/2$, $P(X \le 0) = F_X(0) = 1/2$, and $P(0 < X \le 1) = P(X \le 1) - P(X \le 0) = F_X(1) - 1/2 = 1 - 1/2 = 1/2$. Since the random variable X has an uncountable space (or "range") then X is not a discrete random variable (see Definition 1.6.1). Since the cumulative density function F_X is not continuous then X is not a continuous random variable (see Definition 1.7.1). Therefore, the cumulative density function F_X is a *mixture* of continuous and discrete types. Since th eprobability mass function of a discrete random variable involves summation (or series) and the probability density function of a continuous random variable is an integral, then we cannot find a probability density/mass function for the random variable with cumulative density function F_X.

Revised: 11/1/2019