Continuous Random Variables and Probability Distribution

Definition: The random variable $\bar{X}$ is said to be continuous, with probability density function (pdf) $f(x)$, if, for all $a \leq b$,

$$P(a \leq \bar{X} \leq b) = \int_a^b f(x) \, dx$$

Remarks 1: The graph of $f(x)$ is called the density curve.

The area above the interval $[a, b]$ and under the graph of the pdf $f(x)$.

Remark 2: When $h$ is small and $f(x)$ is smooth,

$$\Pr(x \leq \bar{X} \leq x + h) = \int_{x}^{x+h} f(x) \, dx \approx h \cdot f(x)$$

A discrete probability distribution can be approximated by a continuous probability density.
3%: Obviously, we have
\[ f(x) \geq 0 \]
\[ \int_{-\infty}^{\infty} f(x) \, dx = 1 \]

**Example - Uniform (distribution) density**

**Def.** A continuous rv \( X \) is said to have a uniform distribution on the interval \( [a, b] \) if the pdf of \( X \) is

\[ f(x; a, b) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & \text{otherwise} \end{cases} \]

Then for any \( A < a < \phi < b \),

\[ P(a, \phi) = (\phi - a) \left( \frac{1}{b-a} \right) \]

= depends only on the width \( b-a \) of the interval.

A key rule for densities.

Let \( X \) have pdf \( f(x) \). Then for \( C \subseteq \mathbb{R} \),

\[ P(X \in C) = \int_{X \in C} f(x) \, dx \]
Basic difference between continuous and discrete rvs.
Let \( X \) be a continuous rv with pdf \( f(x) \).

\[
P(\bar{X} = \infty) = \int_{-\infty}^{\infty} f(u) \, du = 0, \quad \forall x \in \mathbb{R}
\]

(Zero probability condition)

\[
P(a \leq \bar{X} \leq b) = \mathbb{P}(a < \bar{X} < b) = \mathbb{P}(a \leq \bar{X} \leq b)
\]

If \( \bar{X} \) is discrete, \( \mathbb{P}(\bar{X} = \infty) \) is not zero for any \( x \) possible values, and all four of these probabilities are different if both \( a \) and \( b \) are possible values.

Example: Exponential density

This most important density arose as an approximation to the geometric distribution with

\[
f(x) = \int_0^\infty \lambda e^{\lambda x} \quad x \geq 0
\]

where \( \lambda > 0 \) is fixed, and \( \lambda \) is up with a (large, \( p \to 0 \))

**Histogram of the geometric distribution** \( \mathbb{P}(k) = \mathbb{P}(1-p)^k p \),

Together with the continuous approximation \( \lambda e^{-\lambda x} \) (broken line).
Clearly we have
\[ \int_{-\infty}^{\infty} f(x) \, dx = 1 \]

* The cumulative distribution function (cdf)

**Definition.** Let \( X \) have pdf \( f(x) \). Then the cdf \( F(x) \) of \( X \) is given by

\[ F(x) = \int_{-\infty}^{x} f(y) \, dy = \mathbb{P}(X \leq x) = \mathbb{P}([X \leq x]) \]

**Remarks:**
- The cdf \( F(x) \) is defined in terms of the pdf \( f(x) \) via integration.
- The pdf \( f(x) \) can be derived from the cdf \( F(x) \) by differentiation.

**Proposition:**

\[ f(x) = \frac{d}{dx} [F(x)] = F'(x) \]

(fundamental theorem of calculus)

**Examples:**
- Uniform distribution

\[ f(x) = (b-a)^{-1} \Rightarrow F(x) = \frac{x-a}{b-a}, \quad x \in [A,B] \]
2\text{/ Exponential distribution.}

\[ f(x) = \lambda e^{-\lambda x} \quad \Rightarrow \quad F(x) = 1 - e^{-\lambda x} \quad \omega \geq 0 \]

\[ F(x) \quad \omega \geq 0 \]

\[ \xrightarrow{x} \]

\[ \therefore \text{Using } F(x) \text{ to compute probability.} \]

\[ P(\overline{x} > a) = 1 - F(a) \]

\[ P(a \leq \overline{x} \leq b) = F(b) - F(a) \]