Chapter 5

Cumulative distribution functions and their applications

Wei-Yang Lin

Department of Computer Science

& Information Engineering

mailto:wylin@cs.ccu.edu.tw

- **5.1 Continuous random variables**

- If X is a continuous random variable with density f, then **cumulative distribution function** (cdf) is defined by

$$F_X(x) := \mathsf{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt. \tag{1}$$

- Pictorially, $F(x)$ is the area under the density $f(t)$ from $-\infty < t \leq x$.

- This is the area of the shaded region in Figure 1.

- Since the total area under a density is one, the area of the unshaded region must be $1 - F(x)$.

- For $a < b$, we can use the cdf to compute probabilities of the form

$$
\begin{aligned}
\mathsf{P}(a \le X \le b) &= \int_a^b f(t)dt \\
&= \int_{-\infty}^b f(t)dt - \int_{-\infty}^a f(t)dt \\
&= F(b) - F(a).
\end{aligned}
$$

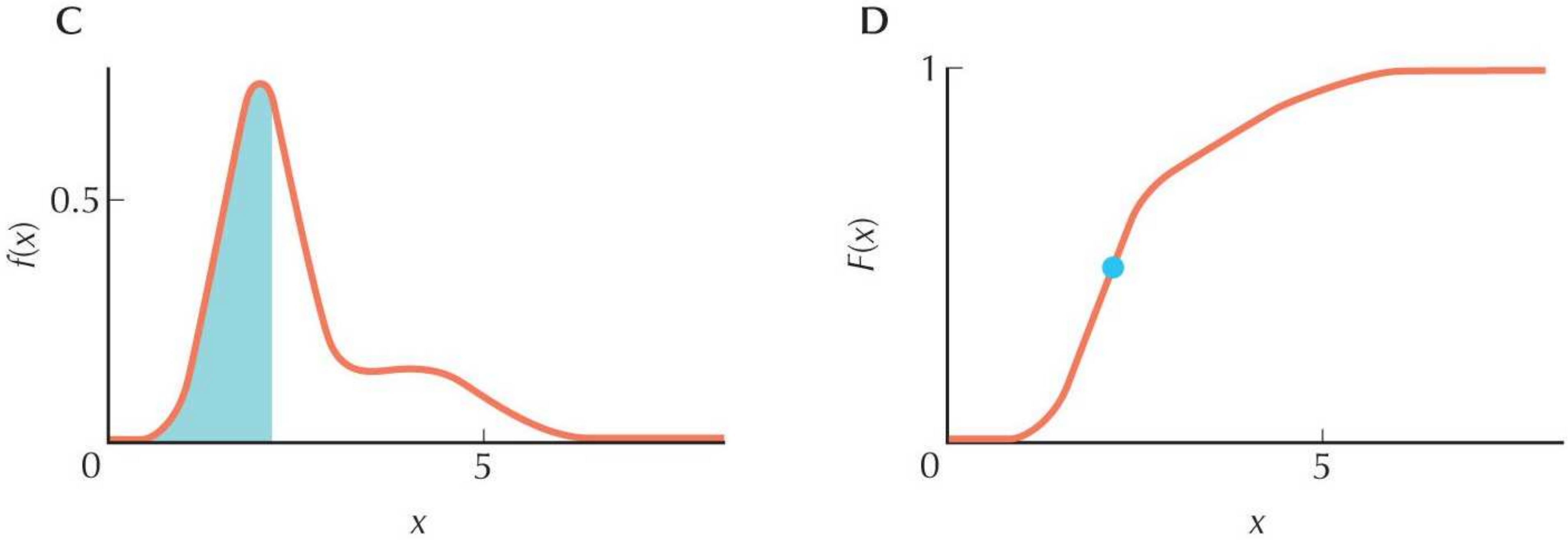

Figure 1: The relationship between pdf $f(x)$ and cdf $F(x)$.

Example 5.2

- Find the cdf of a uniform$[a, b]$ random variable X.

Solution

- Since $f(t) = 0$ for $t < a$, we can see that $F(x) = \int_{-\infty}^{x} f(t)dt$ is equal to 0 for $x < a$.

- For $a \leq x \leq b$, we have

$$F_X(x) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}.$$

- For $x > b$, we have $F_X(x) = 1$.

- We now consider the cdf of a Gaussian random variable.
- If $X \sim N(m, \sigma^2)$, then

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{t-m}{\sigma}\right)^2\right] dt. \tag{2}$$

- Unfortunately, there is no closed-form expression for this integral.
- However, it can be computed numerically.
- In MATLAB, the above integral can be computed with `normcdf(x,m,sigma)`.

- We next show that the $N(m,\sigma^2)$ cdf can always be expressed using the **standard normal cdf**

$$\Phi(y) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\theta^2/2} d\theta, \tag{3}$$

- In (2), make change of variable $\theta = (t-m)/\sigma$ to get

$$\begin{aligned} F_X(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-m)/\sigma} e^{-\theta^2/2} d\theta \\ &= \Phi(\frac{x-m}{\sigma}). \end{aligned}$$

Example 5.3

- At the receiver of a digital communication system, thermal noise in the amplifier sometimes causes an incorrect decision to be made.

- For example, if antipodal signals of energy $\mathcal{E}$ are used, then the bit-error probability can be shown to be $\mathsf{P}(X > \sqrt{\mathcal{E}})$, where $X \sim N(0, \sigma^2)$ represents the noise, and σ^2 is the noise power.

- Express the bit-error probability in terms of the standard normal cdf Φ.

Solution

- The calculation shows that the bit-error probability is completely determined by $\mathcal{E}/\sigma^2$, which is called the **signal-to-noise ratio** (SNR).

$$\mathsf{P}(X > \sqrt{\mathcal{E}}) = 1 - F_X(\sqrt{\mathcal{E}}) = 1 - \Phi\Big(\frac{\sqrt{\mathcal{E}}}{\sigma}\Big)$$

- As the SNR increases, so does Φ, while the error probability decreases.
- In other words, increasing the SNR decreases the error probability.

- For continuous random variables, the density can be recovered from the cdf by differentiation.
- Since
$$F(x) = \int_{-\infty}^{x} f(t)dt,$$
differentiation yields
$$\frac{d}{dx}F(x) = f(x). \tag{4}$$
- Differentiation under the integral sign is described in Wikipedia.

Example 5.5

- Let the random variable X have cdf

$$F_X(x) := \begin{cases} \sqrt{x}, & 0 < x < 1, \\ 1, & x \geq 1, \\ 0, & x \leq 0. \end{cases}$$

- Find the density (pdf).

Solution

- For $0 < x < 1$, $f_X(x) = F'_X(x) = \dfrac{1}{2\sqrt{x}}$.
- For other values of x, $f_X(x) = F'_X(x) = 0$.
- Hence,

$$f_X(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Example 5.6

- Consider an electrical circuit whose random input voltage X is first amplified by $\mu > 0$ and then added a constant offset voltage β.
- If the input is a continuous random variable, find the density of the output.

Solution

- Although the question asks for the density, it is more advantageous to find the cdf first and then obtain the density.
- Let Y denote the output voltage.

$$\begin{aligned} F_Y(y) &= \mathsf{P}(Y \le y) \\ &= \mathsf{P}(\mu X + \beta \le y) \\ &= \mathsf{P}(X \le \frac{y-\beta}{\mu}), \quad \text{since } \mu > 0 \\ &= F_X(\frac{y-\beta}{\mu}). \end{aligned}$$

- If X has density f_X, then

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= \frac{d}{dy} F_X(\frac{y-\beta}{\mu}) \\ &= F_X'(\frac{y-\beta}{\mu})\frac{1}{\mu} \\ &= \frac{1}{\mu} f_X(\frac{y-\beta}{\mu}). \end{aligned}$$

- Recall the chain rule

$$\frac{d}{dy} F(G(y)) = F'(G(y))G'(y).$$

Homework

- Problems 6, 7, 8, 9, 11, 14, 16.