Seoul National University
Dept. of Civil and Environmental Engineering

Instructor: Junho Song
junhosong@snu.ac.kr

# 457.212 Statistics for Civil & Environmental Engineers

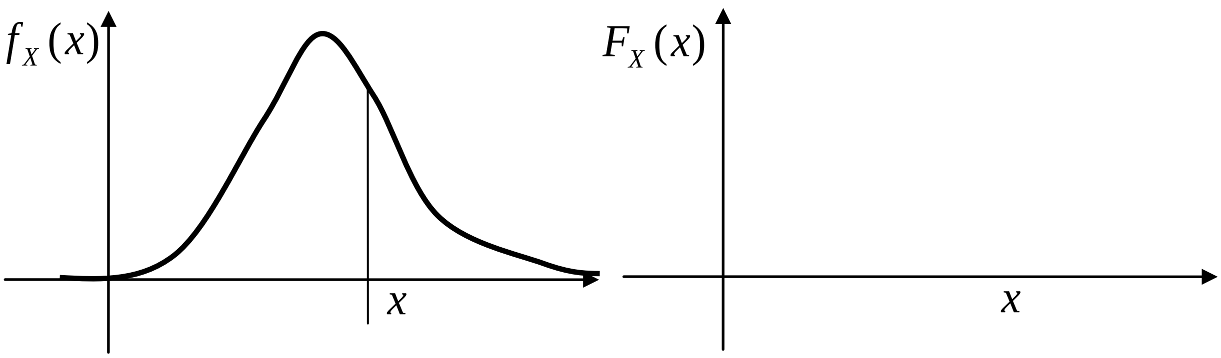
## In-Class Material: Class 10

## Probability Distribution Functions and Partial Descriptors (A&T: 3.1)

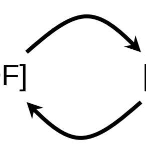
1. **Cumulative Distribution Function** (CDF) of a **continuous** random variable $X$, $F_X(x)$

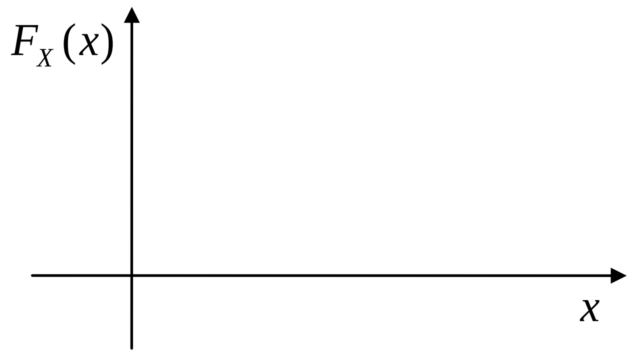
   (a) Definition

   $$F_X(x) = P(\qquad\qquad)$$
   $$= \int f_X(x)dx$$




   (b) Properties

   - $\dfrac{dF_X(x)}{dx} =$ . In summary, [PDF]  [CDF]
   - Non-decreasing (because of integrating non-negative)
   - $F_X(-\infty) = P(\qquad\qquad) =$
   - $F_X(\infty) = P(\qquad\qquad) =$



Instructor: Junho Song
junhosong@snu.ac.kr

**Example 1:** Suppose the likelihood of accidents is uniform along the 100 km highway. Let $X$ denote the distance between the starting point and an accident location. Determine

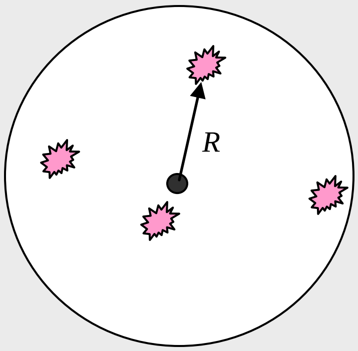

(a) Probability density function (PDF) of $X$ and plot

(b) Cumulative density function (CDF) of $X$ and plot:

(c) $P(20 \le X \le 50)$ by use of PDF and CDF


Instructor: Junho Song
junhosong@snu.ac.kr

**Example 2:** Suppose an object can fall anywhere within a 100-km radius circle at random (i.e. uniform likelihood over points inside the circle). PDF and CDF of the distance between the location and the center of the circle, $R$ ?




2. **Partial Descriptors** of a random variable

(a) "Complete" description by probability functions:

(b) "Partial" descriptors: measures of key characteristics; can derive from ( )

**Note:**

- Expectation: $E[\cdot] = \int_{-\infty}^{\infty} (\cdot) f_X(x) dx$ (continuous) or $\sum_{\text{all } x} (\cdot) p_X(x)$ (discrete)
- Moment: $E[\cdot X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$ or $\sum_{\text{all } x} x^n p_X(x)$
- Central Moment, $E[(X - \mu_X)^n] = \int_{-\infty}^{\infty} (x - \mu_X)^n f_X(x) dx$ or $\sum_{\text{all } x} (x - \mu_X)^n p_X(x)$

Instructor: Junho Song
junhosong@snu.ac.kr

| | Name | Definition | Meaning (PDF/CDF) |
|---|---|---|---|
| Measure of Central Location | Mean, $\mu_X$ | First moment, $\mathrm{E}[X]$ | Location of the ( ) of an area underneath ( ) |
| | Median, $x_{0.5}$ | $F_X(x_{0.5}) = 0.5$<br>$F_X^{-1}(0.5)$ | The value of a r.v. at which values above and below it are _______lly probable.<br>If symmetric? |
| | Mode, $\tilde{x}$ | $\arg\max_x f_X(x)$ | The outcome that has the _______est probability mass or density |
| Measure of Dispersion | Variance, $\sigma_X^2$ | Second-order central moment<br>$\mathrm{E}[(X-\mu_X)^2]$<br>$=\mathrm{E}[X^2]-\mathrm{E}[X]^2$ | Average of squared deviations |
| | Standard Deviation, $\sigma_X$ | $\sqrt{\sigma_X^2}$ | Radius of ( ) |
| | Coefficient of Variation (C.O.V.), $\delta_X$ | $\dfrac{\sigma_X}{\lvert\mu_X\rvert}$ | __________ed radius of ( ) |
| Asymmetry | Coefficient of Skewness, $\gamma_X$ | Third-order central moment normalized by $\sigma_X^3$,<br>$\dfrac{\mathrm{E}[(X-\mu_X)^3]}{\sigma_X^3}$ | Behavior of two tails<br>$> 0$<br>$= 0$<br>$< 0$ |
| Flatness | Coefficient of Kurtosis, $\kappa_X$ | Fourth-order central moment normalized by $\sigma_X^4$,<br>$\dfrac{\mathrm{E}[(X-\mu_X)^4]}{\sigma_X^4}$ | "Peakedness" - more of the variance is due to infrequent extreme deviations, as opposed to frequent modestly-sized deviations. |

Instructor: Junho Song
junhosong@snu.ac.kr

**Example 3:** Compute the mean, median, mode, variance, standard deviation and coefficient of variation for a discrete random variable $X$ whose PMF is given as

| $x$ | $P_X(x)$ |
|---|---|
| 0 | 0.20 |
| 1 | 0.50 |
| 2 | 0.30 |

**Example 4:** Compute the mean, median, mode, variance, standard deviation and coefficient of variation for a continuous random variable $X$ whose PDF is given by $f_X(x) = (3/1000)x^2, \quad 0 \le x \le 10$ and 0 elsewhere. What is your guess on the sign of the coefficient of the skewness? Confirm your guess by computing it.