CMPE 320: Probability, Statistics, and Random Processes

Lecture 14: Cumulative distribution functions

Spring 2018

Seung-Jun Kim

Cumulative distribution function (CDF)

- CDF "accumulates" probability "up to" the value x

$$\underbrace{F_X(x)}_{\text{notation for CDF}} = P(X \le x)$$

- CDF can describe both discrete and continuous RVs

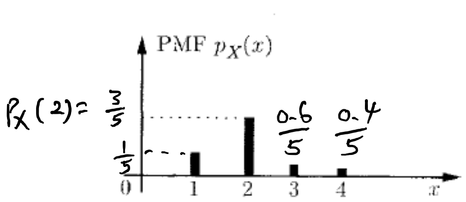
Discrete RV : $F_X(x) = P(X \le x) = \sum_{k \le x} P_r(X = k) = \sum_{k \le x} P_X(k)$

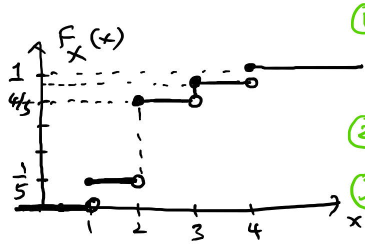
Continuous RV : $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)\, dt$

- Any specification of probability of events $\{X \le x\}$ (be it through PMF, PDF, or CDF) is probability law of RV X

PMF to CDF

$$F_X(x) = \sum_{k \le x} p_X(k)$$

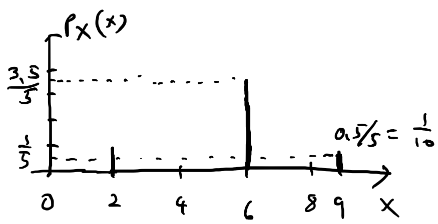


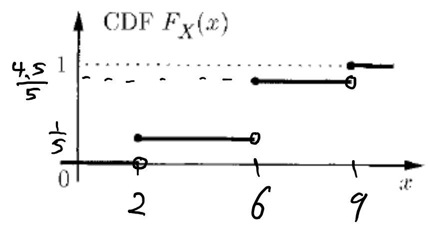


① CDF must be monotonically nondecreasing

② $F_X(-\infty) = 0$

③ $F_X(\infty) = 1$

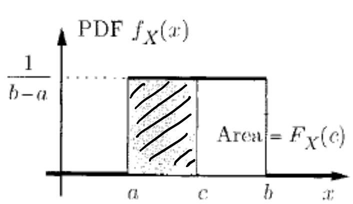


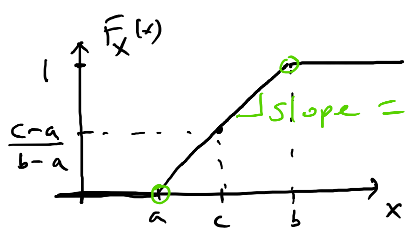


$$p_X(k) = F_X(k) - F_X(k-1)$$

PDF to CDF

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$





slope $= \frac{\Delta y}{dx} = \frac{1-0}{b-a} = \frac{1}{b-a}$

Alternatively

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

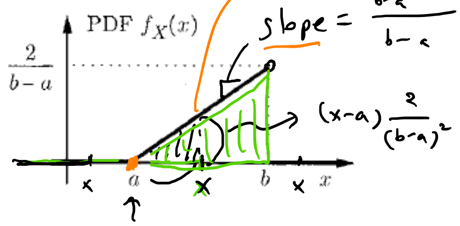
$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \begin{cases} 0, & \text{if } x < a \\ \int_a^x \frac{1}{b-a}\, dt = \frac{t}{b-a}\Big|_a^x = \frac{x-a}{b-a}, & a \le x \le b \\ 1, & \text{if } x > b \end{cases}$$

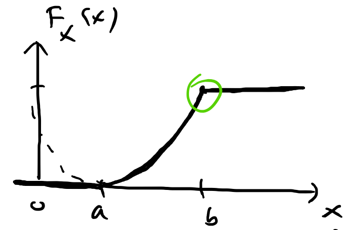
PDF to CDF (2)

$(x_0, y_0) = (a, 0)$

$y - y_0 = m(x - x_0) \Rightarrow y - 0 = \frac{2}{(b-a)^2}(x-a)$

slope $= \dfrac{\frac{2}{b-a} - 0}{b-a} = \dfrac{2}{(b-a)^2}$





If $x < a$ $\quad F_X(x) = 0$

If $a \le x \le b$ $\quad F_X(x) = \frac{1}{2}(x-a)\cdot(x-a)\frac{2}{(b-a)^2} = \frac{(x-a)^2}{(b-a)^2}$

If $x > b$ $\quad F_X(x) = 1$

$$f_X(x) = \begin{cases} 0 & x < a,\ x > b \\ \frac{2(x-a)}{(b-a)^2}, & a \le x \le b \end{cases}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0 & , x < a \\ \int_a^x \frac{2(t-a)}{(b-a)^2}dt = \frac{2}{(b-a)^2}\int_a^x (t-a)dt = & , a \le x \le b \\ 1 & x > b \end{cases}$$

$\frac{2}{(b-a)^2}\left[\frac{1}{2}t^2 - at\right]_a^x$

$= \frac{2}{(b-a)^2}\left(\frac{1}{2}x^2 - ax - \left(\frac{1}{2}a^2 - a^2\right)\right)$

$= \frac{1}{(b-a)^2}(x^2 - 2ax + a^2)$

$= \frac{(x-a)^2}{(b-a)^2}$

Properties of CDF

1. $F_X(x)$ is monotonically non-decreasing

   that is, if $x < y \Rightarrow F_X(x) \le F_X(y)$

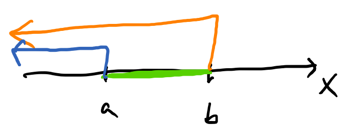
2. $F_X(-\infty) = 0$

3. $F_X(\infty) = 1$

4. $P(a \le X \le b) = P(X \le b) - P(X \le a)$

   $= F_X(b) - F_X(a)$

$F_X(a) = P(X \le a)$



PMF or PDF from CDF

- Discrete RV X

If X takes integer values,

$$\underline{p_X(k) = F_X(k) - F_X(k-1)}$$

- Continuous RV X

$$\boxed{f_X(x) = \frac{d}{dx} F_X(x)}$$

UMBC CMPE 320 Seung-Jun Kim

Example 3.6. The Maximum of Several Random Variables. You are allowed to take a certain test three times, and your final score will be the maximum of the test scores. Thus,

$$X = \max\{X_1, X_2, X_3\},$$

where X_1, X_2, X_3 are the three test scores and X is the final score. Assume that your score in each test takes one of the values from 1 to 10 with equal probability 1/10, independently of the scores in other tests. What is the PMF p_X of the final score?

First compute $F_X(x)$ and then $P_X(k) = F_X(k) - F_X(k-1)$

$$F_X(x) = P(X \le x) = P(\max\{X_1, X_2, X_3\} \le x) = P(X_1 \le x \cap X_2 \le x \cap X_3 \le x)$$

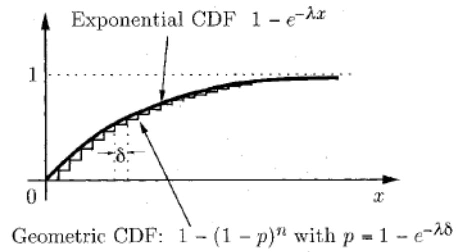
$$= P(X_1 \le x) P(X_2 \le x) P(X_3 \le x) = \underbrace{F_{X_1}(x)} F_{X_2}(x) F_{X_3}(x)$$

$$P_{X_i}(x) = \begin{cases} \frac{1}{10}, & \text{if } 1 \le x \le 10 \\ 0, & \text{otherwise} \end{cases} \Rightarrow F_{X_i}(x) = \sum_{k \le x} P_{X_i}(k) = \begin{cases} 0, & x < 1 \\ \frac{x}{10}, & 1 \le x \le 10 \\ 1 & x > 10 \end{cases}$$

$$F_X(x) = \begin{cases} 0, & x < 1 \\ \left(\frac{x}{10}\right)^3 & 1 \le x \le 10 \\ 1 & x > 10 \end{cases} \Rightarrow P_X(k) = F_X(k) - F_X(k-1) = \begin{cases} \left(\frac{k}{10}\right)^3 - \left(\frac{k-1}{10}\right)^3, & 1 \le x \le 10 \\ 0, & \text{otherwise} \end{cases}$$

Geometric and exponential CDFs

- Find the CDF of a geometric RV X. Repeat with an exponential RV X.



Problem 6. Calamity Jane goes to the bank to make a withdrawal, and is equally likely to find 0 or 1 customers ahead of her. The service time of the customer ahead, if present, is exponentially distributed with parameter λ. What is the CDF of Jane's waiting time?

If X is exponential with parameter λ, $f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$

W: Jane's waiting time $\quad F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0 & x < 0 \end{cases}$

$$F_W(w) = P(W \leq w) = P(W \leq w \cap 0 \text{ customer}) + P(W \leq w \cap 1 \text{ customer})$$

$$= P(W \leq w \mid 0 \text{ customer}) P(0 \text{ customer}) + P(W \leq w \mid 1 \text{ customer}) P(1 \text{ customer})$$

$$= \left(\begin{cases} 0 & \text{if } w < 0 \\ 1 & \text{if } w \geq 0 \end{cases}\right) \cdot \frac{1}{2} + \left(\begin{cases} 1 - e^{-\lambda w}, & \text{if } w \geq 0 \\ 0 & \text{if } w < 0 \end{cases}\right) \frac{1}{2}$$

$$= \begin{cases} 0 & \text{if } w < 0 \\ \frac{1}{2} + \frac{1}{2}(1 - e^{-\lambda w}) & \text{if } w \geq 0 \end{cases}$$