Lecture 11

Continuous Random Variables

Manju M. Johny

STAT 330 - Iowa State University

Continuous Random Variables

Discrete vs. Continuous R.Vs

Discrete Random Variable	Continuous Random Variable
Sample space (Ω) maps to finite or countably infinite set in $\Re$	Sample space (Ω) maps to an uncountable set in $\Re$.
Ex: $\{1, 2, 3\}$, $\{1, 2, 3, 4, \ldots\}$	Ex: $(0, \infty)$, $(10, 20)$

- We have already learned about discrete R.Vs (Lectures 5-10)
- All properties of discrete R.Vs have direct counterparts for continuous R.Vs
- Summations (Σ) used for discrete R.V's are replaced by integrals ($\int$) for continuous R.V's.

Discrete X

$$E X = \sum_x x \, P_X(x)$$

Continuous X

$$E X = \int x \, f(x) \, dx$$

CDF of Continuous Random Variables

Definition

Let X be a continuous random variable. The *cumulative distribution function (cdf)* of X is

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

CDF

$F_X(t)$

- All cdf properties discussed earlier still hold
  1. $0 \le F_X(t) \le 1$
  2. F_X is non-decreasing (if $a \le b$, then $F_X(a) \le F_X(b)$.
  3. $\lim_{t \to -\infty} F_X(t) = 0$ and $\lim_{t \to \infty} F_X(t) = 1$
  4. F_X is right-continuous with respect to t

- The cdf for continuous R.V is also continuous (not a step function like in discrete case)
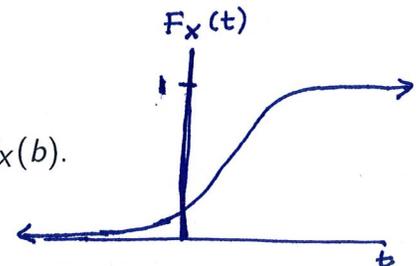
Definition

For a continuous variable X with cdf F_X, the *probability density function (pdf)* of X is defined as:

$$f(x) = F_X'(x) = \frac{d}{dx}F_X(x)$$

Properties of pdf:

1. $f(x) \geq 0$ for all x,
2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

PDF

$f(x)$
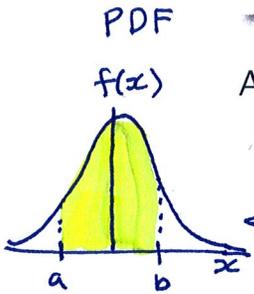
Additionally, for continuous R.V X,

- $F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f(x)dx$ for any $t \in \mathbb{R}$
- $P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$ for any $a, b \in \mathbb{R}$
- $P(X = a) = P(a \leq X \leq a) = \int_{a}^{a} f(x)dx = 0$ for any $a \in \mathbb{R}$

$f(x)$ PDF

(The probability of X being any single) value is 0.

$\rightarrow P(x \leq t) = P(x < t)$
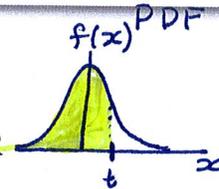
Examples

Example 1: Let Y be the time (in yrs) until the first major failure of a new disk drive. Suppose the probability density function (pdf) $\underline{\text{PDF}}$ of Y is given by

$$f(y) = \begin{cases} 0 & y \leq 0 \\ e^{-y} & y > 0 \end{cases}$$

PDF

$f(y)$
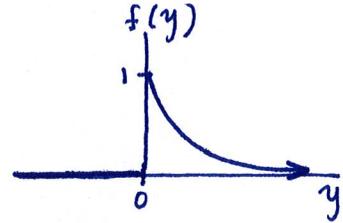
1. Check whether $f(y)$ is a *valid* density function.

   We need to check the 2 properties of pdfs.

   (1) $f(y)$ is non-negative function on $\mathbb{R}$ ✓

   (2) $\int_{-\infty}^{\infty} f(y)dy = 1$ ✓

$$\int_{-\infty}^{\infty} f(y)dy = \int_{-\infty}^{0} 0\,dy + \int_{0}^{\infty} e^{-y}\,dy = -e^{-y}\Big|_{0}^{\infty} = 0 - -e^{0}$$

$$= 1$$

2. What is the probability that the 1st major disk drive failure occurs within the first year?

PDF

$f(y)$

e^{-y}
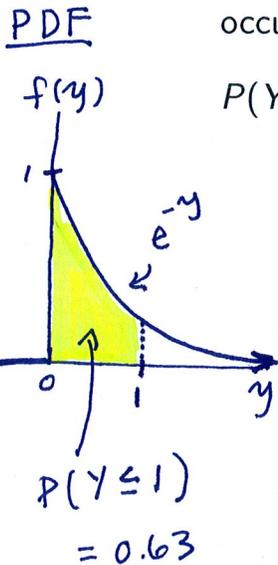
$$P(Y \leq 1) = \int_{-\infty}^{1} f(y)\,dy$$

$$= \int_{-\infty}^{0} 0\,dy + \int_{0}^{1} e^{-y}\,dy$$

$$= -e^{-y}\Big|_{0}^{1}$$

$P(Y \leq 1)$

$= 0.63$

$$= -e^{-1} - -e^{0}$$

$$= -e^{-1} + 1$$

$$= 1 - e^{-1}$$

$$= 0.63$$

3. What is the probability that the 1st major disk drive failure occurs before the first year?

$f(y)$

$$P(Y < 1) = \int_{-\infty}^{1} f(y)dy$$

$$= \int_{0}^{1} e^{-y}dy$$

$$= -e^{-y}\Big|_{0}^{1}$$

$$= 1 - e^{-1}$$

$$= 0.63$$

(This is same as $P(Y \leq 1)$ from prev. slide

$$\left[\begin{array}{c} \text{For continuous dist,} \\ P(Y \leq a) = P(Y < a) \\ b/c \ P(Y = a) = 0 \end{array}\right]$$

PDF

$f(y)$
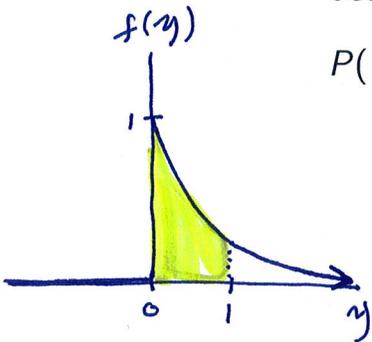
4. What is the probability that the 1st major disk drive failure occurs after the first year?

$$P(Y > 1) = \int_{1}^{\infty} f(y)dy$$

$$= \int_{1}^{\infty} e^{-y}dy$$

$$= -e^{-y}\Big|_{1}^{\infty}$$

$$= 0 - -e^{-1}$$

$$= e^{-1}$$

$$= 0.37$$

We can also get this by

$P(Y > 1)$
$= 1 - P(Y \leq 1)$
$= 1 - 0.63$
$= 0.37$
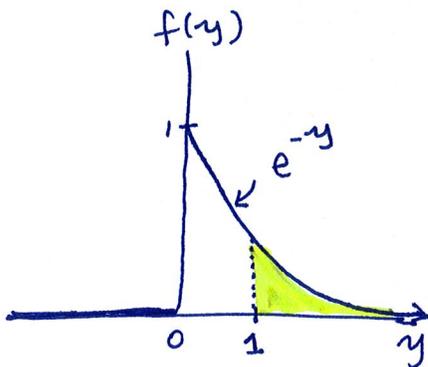
5. What is the probability that the 1st major disk drive failure occurs after first year but before second year?

PDF

$f(y)$
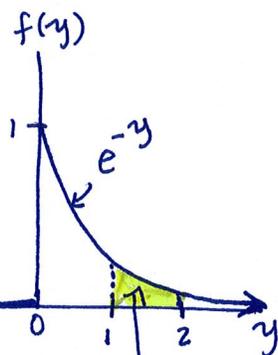
$P(1 < Y < 2)$
$= 0.23$

$$P(1 < Y < 2) = \int_1^2 f(y)\,dy$$

$$= \int_1^2 e^{-y}\,dy$$

$$= -e^{-y}\Big|_1^2$$

$$= (-e^{-2}) - (-e^{-1})$$

$$= -e^{-2} + e^{-1}$$

$$= 0.23$$

6. What is the cumulative distribution function (cdf) of Y?

$f(y)$

① For $t \leq 0$

$$F_Y(t) = P(Y \leq t) = \int_{-\infty}^t f(y)\,dy = \int_{-\infty}^t 0\,dy = \boxed{0}$$
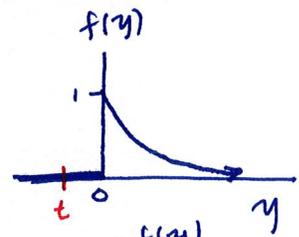
② For $t > 0$

$$F_Y(t) = P(Y \leq t) = \int_{-\infty}^t f(y)\,dy = \int_{-\infty}^0 0\,dy + \int_0^t e^{-y}\,dy$$


$$= -e^{-y}\Big|_0^t = -e^{-t} - -e^{0} = \boxed{1 - e^{-t}}$$
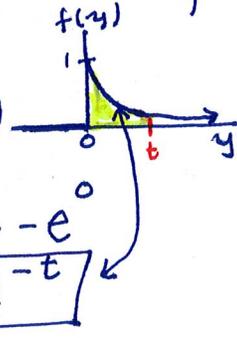
$\boxed{CDF}$

$$F_Y(t) = \begin{cases} 0 & \text{for } t \leq 0 \\ 1 - e^{-t} & t > 0 \end{cases}$$

For Example 1, the pdf and cdf of Y are shown below.

f(y)   Probability Density Function (pdf) of Y

F(y)   Cumulative Distribution Function (cdf) of Y

Easier b/c we don't have to integrate every time we calculate probabilitie.

Just integrate once to get CDF & plug in values
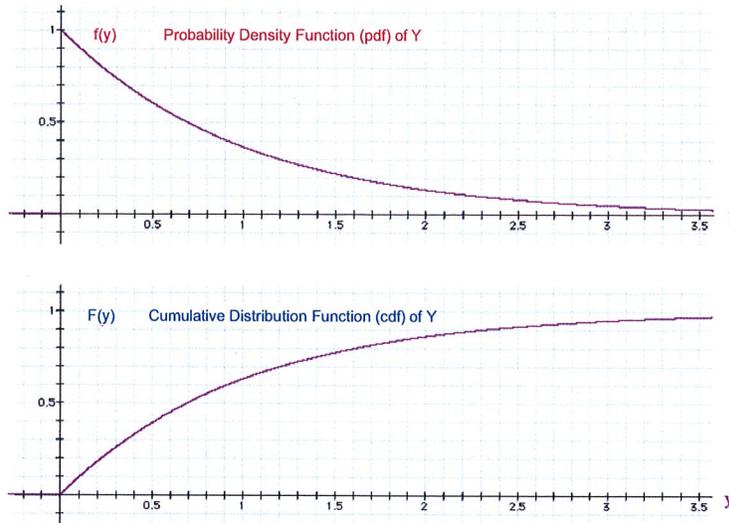
SHORT CUT: Use the cdf to calculate desired probabilities instead of integrating the pdf for each problem.

- Only need to integrate the pdf once to obtain the cdf
- Write any probability in terms of the cdf and plug in to solve

Back to Example 1:   CDF: $F_Y(t) = P(Y \le t) = \begin{cases} 1 - e^{-t} & \text{for } t > 0 \\ 0 & \text{for } t \le 0 \end{cases}$

- $P(Y \le 1) = F_Y(1) = 1 - e^{-1} = 0.63$
  $= P(Y < 1)$

- $P(Y > 1) = 1 - P(Y \le 1) = 1 - F_Y(1) = 1 - 0.63 = 0.37$

- $P(1 < Y < 2) = P(Y < 2) - P(Y \le 1)$
  $= F_Y(2) - F_Y(1)$
  $= (1 - e^{-2}) - (1 - e^{-1})$
  $= -e^{-2} + e^{-1}$
  $= 0.23$

**Example 2:** Let X be a random variable with the following probability density function (pmf):

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2} & \text{for } 0 \leq x \leq \frac{1}{2} \\ 2x & \text{for } \frac{1}{2} < x < 1 \\ 0 & \text{for } x \geq 1 \end{cases}$$

can also be written as

$$f(x) = \begin{cases} \frac{1}{2} & 0 \leq x \leq \frac{1}{2} \\ 2x & \frac{1}{2} < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

we have 4 sections

PDF
$f(x)$

$f(x) = \frac{1}{2}$
$f(x) = 2x$
$f(x) = 0$
$f(x) = 0$
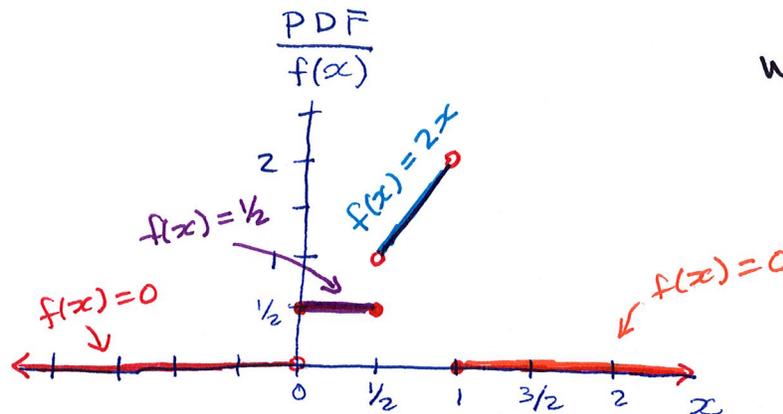
13 / 18

1. Give the cumulative distribution function (cdf) of X $\left(\begin{array}{c}\text{Find area} \\ \text{under PDF curve}\end{array}\right)$

$$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f(x)\, dx$$

Consider 4 sections seperately

① If $t < 0$

$$F_X(t) = \int_{-\infty}^{t} f(x)\, dx = \int_{-\infty}^{t} 0\, dx \boxed{= 0}$$
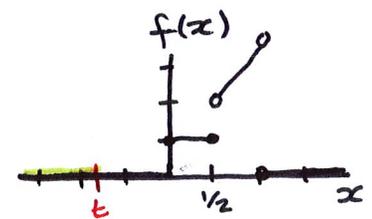
② If $0 \leq t \leq \frac{1}{2}$

$$F_X(t) = \int_{-\infty}^{t} f(x)\, dx$$

$$= \int_{-\infty}^{0} 0\, dx + \int_{0}^{t} \frac{1}{2}\, dx$$

$$= \frac{x}{2}\Big|_{0}^{t} = \frac{t}{2} - \frac{0}{2} \boxed{= \frac{t}{2}}$$
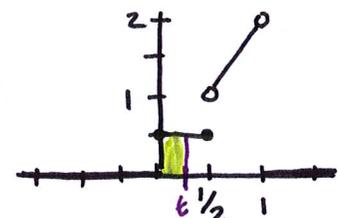

14 / 18

$\frac{1}{2} < t \leq 1$

③ $F_X(t) = \int_{-\infty}^{t} f(x)\,dx$

$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{1/2} \frac{1}{2}\,dx + \int_{1/2}^{t} 2x\,dx$

$= \frac{x}{2}\Big|_{0}^{1/2} + x^2\Big|_{1/2}^{t}$

$= \left(\frac{1/2}{2} - \frac{0}{2}\right) + \left(t^2 - (1/2)^2\right) = \frac{1}{4} + t^2 - \frac{1}{4} \boxed{= t^2}$
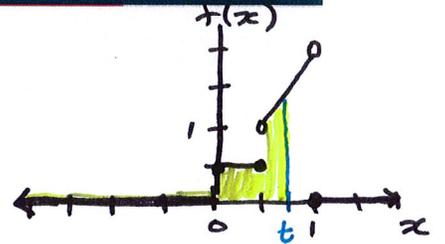
④ $t \geq 1$

$F_X(t) = \int_{-\infty}^{t} f(x)\,dx$

$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{1/2} \frac{1}{2}\,dx + \int_{1/2}^{1} 2x\,dx + \int_{1}^{t} 0\,dx$

$= \frac{x}{2}\Big|_{0}^{1/2} + x^2\Big|_{1/2}^{1} = \left(\frac{1/2}{2} - \frac{0}{2}\right) + \left(1^2 - (1/2)^2\right) = \frac{1}{4} + 1 - \frac{1}{4} \boxed{= 1}$
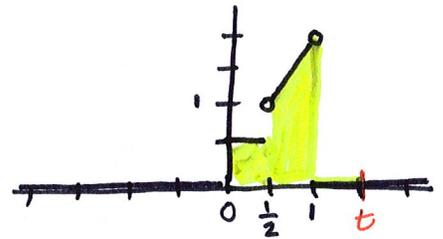
So the cdf of X is . . . .

CDF
$F_X(t)$

$F_X(t) = \begin{cases} 0 & \text{if } t < 0 \\ t/2 & \text{if } 0 \leq t \leq 1/2 \\ t^2 & \text{if } 1/2 < t < 1 \\ 1 & \text{if } t \geq 1 \end{cases}$

$F_X(1/2) = 1/4$

16 / 18
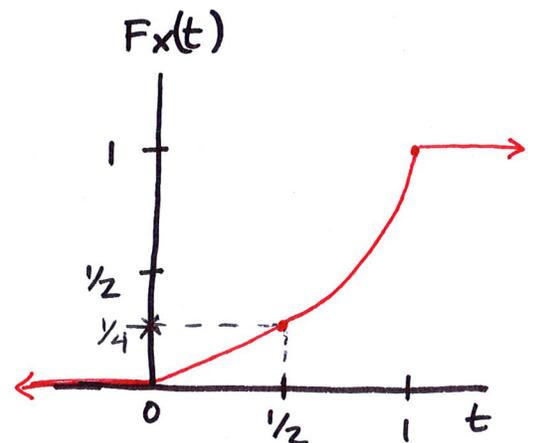
2. What is the probability that X is less that 0.75? plug into cdf ↙ directly

$$P(X < 0.75) = F_X(0.75) = (0.75)^2 = 0.5625$$

otherwise, I have to integrate pdf

$$P(X < 0.75) = \int_{-\infty}^{0.75} f(x)\,dx$$

$f(x)$
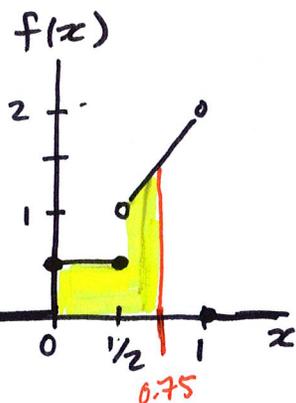
$$= \int_{-\infty}^{0} 0\,dx + \int_{0}^{1/2} \tfrac{1}{2}\,dx + \int_{1/2}^{0.75} 2x\,dx$$

$$= \left.\frac{x}{2}\right|_{0}^{1/2} + \left.x^2\right|_{1/2}^{0.75}$$

$$= \frac{1}{4} + \left[0.75^2 - (1/2)^2\right]$$

$$= 0.75^2 = 0.5625$$

Summary of Discrete & Continuous R.V.

Discrete R.V.

- Im(X) finite or countable infinite

- CDF: $F_X(t) = P(X \le t)$
  $$= \sum_{x \le t} p_X(x)$$

- PMF: $p_X(x) = P(X = x)$

- $E(h(X)) = \sum_x h(x) p_X(x)$

- $E(X) = \sum_x x p_X(x)$

- $Var(X) = E(X^2) - [E(X)]^2$

Continuous R.V.

- Im(X) uncountable

- CDF: $F_X(t) = P(X \le t)$
  $$= \int_{-\infty}^{t} f(x)\,dx$$

- PDF: $f_X(x) = \frac{d}{dx} F_X(x)$

- $E(h(X)) = \int_x h(x) f(x)\,dx$

- $E(X) = \int_{-\infty}^{\infty} x f(x)\,dx$

- $Var(X) = E(X^2) - [E(X)]^2$