# STAT 320: Principles of Probability Unit 6 Part:A Continuous Random Variables

United Arab Emirates University

Department of Statistics

### Outline

Characterization of any CDF function

Reminder: The Cumulative **Distribution Functions** 

### Distribution Functions

### Definition (Cumulative Distribution Function (cdf))

The **cumulative distribution function** or **cdf** of a *any* variable X, denoted by  $F_{x}(x)$ , is defined by

$$F_X(x) = P(X \le x)$$
 for all  $x \in \mathbb{R}$ .

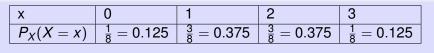
# CDF: Example

Consider the experiment of tossing three fair coins, and let X = number of heads observed. We have already seen that

Х	0	1	2	3
$p_{\chi}(x)$	1 8	3 8	3 8	1 8

#### The cdf of X is:

$$F_{x}(x) = \begin{cases} 0 & \text{if } -\infty < x < 0 \\ \frac{1}{8} & \text{if } 0 \le x < 1 \\ \frac{4}{8} & \text{if } 1 \le x < 2 \\ \frac{7}{8} & \text{if } 2 \le x < 3 \\ 1 & \text{if } 3 \le x < \infty \end{cases}$$



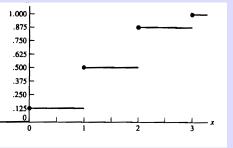


Figure: The polt of  $F_X(x)$ : CDF of the random variable X

Note that  $F_X(\cdot)$  is defined for all values of  $x \in \mathbb{R}$ , not just for  $x \in \mathbb{S}_x :=$  $\{0,1,2,3\}$ . For example,  $2.5 \notin \mathbb{S}_{r}$ , however

$$F_X(2.5) = P_X(x \le 2.5) = P_X(X = 0) + P_X(X = 1) + P_X(X = 2) = \frac{7}{8}.$$



### Characterization of a CDF

#### Theorem

The function F(x) is a cdf if and only if the following three conditions hold:

- $\lim_{x \to -\infty} F(x) = 0 \text{ and } \lim_{x \to \infty} F(x) = 1.$
- F(x) is a nondecreasing function of x
- F(x) is right-continuous; that is, for every real number  $x_0$ ,  $\lim F(x) = F(x_0).$  $X \setminus X_0$

Comment: Let X be a random variable with the corresponding cdf  $F_X(x)$  for  $x \in \mathbb{R}$ . Let  $x_0 \in \mathbb{R}$  is arbitrary. Then

$$P(X=x_0):=P(X\in\{x_0\})=\lim_{X\ \downarrow\ X_0}F_X(X)-\lim_{X\ \uparrow\ X_0}F_X(X).$$

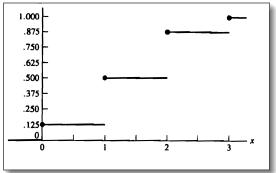


Figure: The polt of  $F_X(x)$ : CDF of the random variable X

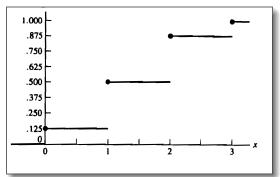


Figure: The polt of  $F_X(x)$ : CDF of the random variable X

Let  $F_X(x)$  denotes the cdf function included in the above image. Therefore.

$$P(X = 1) = \lim_{x \downarrow 1} F_X(X) - \lim_{x \uparrow 1} F_X(x) = 0.5 - 0.125 = 0.375.$$

$$P(X = 1.5) = \lim_{x \downarrow 1.5} F_X(X) - \lim_{x \uparrow 1.5} F_X(x) = 0.5 - 0.5 = 0.$$

### **Example:** An example of a continuous cdf is the function

$$F_X(x) := \frac{1}{1 + e^{-x}}$$
 for all  $x \in \mathbb{R}$ .

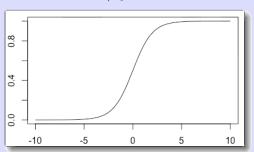


Figure: The polt of  $F_X(x)$ : CDF of the random variable X

Verify: The above function satisfies the three conditions required to be a CDF.

Question: Prove that the following functions are valid cdfs.

- $P(x) = e^{-e^{-x}} \text{ for all } x \in \mathbb{R}.$

### Reminder: Discrete Random Variable

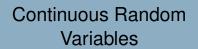
#### Definition (Discrete Random Variable)

A random variable X is discrete if it's support  $\mathbb{S}_{\nu}$  is finite or countable infinite.

Alternative Characterization of Discrete Distributions: A random variable X is discrete if the corresponding cdf  $F_X(x)$  is a step function of x. i.e.  $F_X(x)$  increases only via jumps.

### Outline

Continuous Random Variables



### Continuous and Discrete Random variable

### Definition (Continuous Random Variable)

A random variable X is continuous if the corresponding cumulative distribution function (cdf)  $F_{\nu}(x)$  is a continuous and continuously differentiable (at all points except for a finitely/countably many points) function of x.

### Definition (Probability Density Function)

Let  $F_{\nu}(x)$  be a cumulative distribution function (CDF) of a continuous random variable, then the corresponding probability density function or **pdf**, denoted as  $f_x(x)$  is a function that satisfies the following creteria.

$$F_{\chi}(\mathbf{x}) = \int_{-\infty}^{\mathbf{X}} f_{\chi}(y) dy$$

### Definition (Support of a Continuous Random Variable)

Let X be a continuous random varibale with probability density function  $f_{\nu}(x)$ . The support of the random variable is defined to be

$$\mathbb{S}_{x}=\{x:f_{x}(x)>0\}$$

Support of a random variable refers to the all possible value or outcome of the random variable. Technically, if a point is a possible outcome, the corresponding value of the pdf is positive.

### Relation between pdf and CDF of a continuous random variable

$$F_{X}(\mathbf{x}) = \int_{-\infty}^{\mathbf{X}} f_{X}(y) dy$$

$$f_{X}(x) = \frac{d}{dx}F_{X}(x)$$

# Probability Density Function (pdf): For continuous RV

**Comment:** Using the Fundamental Theorem of Calculus, if  $f_{\nu}(x)$ 

is continuous, we have the further relationship  $f_{\chi}(x) = \frac{d}{dx} F_{\chi}(x)$ .

If X is a continuous random variable, then probabilities can be obtained by integrating its pdf over suitable region. Specifically, for  $a, b \in \mathbb{R}$ , a < b,

$$P\left(\mathbf{a} < X \leq \mathbf{b}\right) = F_{X}\left(\mathbf{b}\right) - F_{X}\left(\mathbf{a}\right) = \int_{\mathbf{a}}^{\mathbf{b}} f_{X}(x)dx.$$

Question: Is it ture that a random variable must be continuous if its support is an interval?

Question: Is it ture that a random variable must be continuous if its support is  $\mathbb{R}$ ?

Question: Is it possible for a continuous random variable to have a support that has finitely many points?

### **Example:** An example of a continuous cdf is the function

$$F_X(x) := \frac{1}{1 + e^{-x}}$$
 for all  $x \in \mathbb{R}$ .

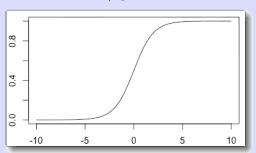


Figure: The polt of  $F_X(x)$ : CDF of the random variable X

**Verify:** The above function satisfies the three conditions required to be a CDF.

### **Example of CDF of a Continuous Random Variable:**

$$F_X(x) := \left\{ \begin{array}{ll} 0 & \text{if } x \leq 0 \\ 1 - e^{-x} & \text{if } x > 0 \end{array} \right.$$

**Verify:** The above function satisfies the three conditions required to be a CDF.

# Characterization of Probability Density Function (pdf)

### Definition (Continuous Random Variable)

A random variable X is said to be continuous if there is a function f(x), called the probability density function (pdf), such that

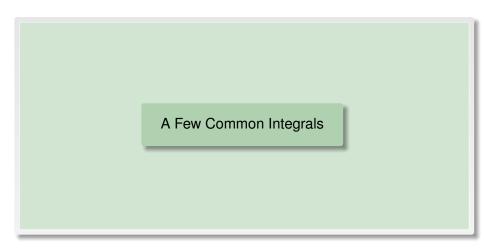
- $f(x) \geq 0$  for all x.
- $P(a \le X \le b) = \int_a^b f(x) dx for all a < b.$

### Result:

If X is a continuous random variable then,

$$P(X=c)=0$$
 for any  $c\in\mathbb{R}$ 

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$$



$$\int x^n dx = \frac{x^{n+1}}{n+1} \text{ for any integer } n, n \neq -1.$$

$$\int \frac{1}{x} dx = \log(x)$$

$$\int e^{-x} dx = -e^{-x}.$$

$$\int e^{mx} dx = rac{e^{mx}}{m}$$
 for any nonzero real number  $m \in \mathbb{R}, \, m 
eq 0.$ 

\* Note: We have not included the constant term that appears as a constants while writing a indefinite integral. For the majority,

if not all, of the integras in this course will be a definite integrals with a lower and upper limit.

Assume 
$$f'(x) := \frac{d}{dx} \frac{f(x)}{dx}$$
 and  $g'(x) := \frac{d}{dx} \frac{g(x)}{dx}$  for the following formula

Integral By Parts: 
$$\int f(x)g(x)dx = f(x) \left( \int g(x)dx \right) - \int \left\{ f'(x) \left( \int g(x)dx \right) \right\} dx$$

Addition Rule: 
$$\int \left\{ c_1 \frac{f(x)}{f(x)} + c_2 \frac{g(x)}{g(x)} \right\} dx = c_1 \int f(x) dx + c_2 \int g(x) dx \text{ for any constant } c_1, c_2 \in \mathbb{R}.$$

Example: Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) := \begin{cases} C(4x - 2x^2) & \text{if } 0 < x < 2\\ 0 & \text{otherwise.} \end{cases}$$

- What is the value of C?
- Find P(X > 1).

Example:

Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) := \begin{cases} C(4x - 2x^2) & \text{if } 0 < x < 2\\ 0 & \text{otherwise.} \end{cases}$$

What is the value of C? Find P(X > 1).

According to the property of the pdf

$$\int f(x)dx = 1$$

$$\Rightarrow \int_0^2 C(4x - 2x^2)dx = 1$$

$$\Rightarrow C(2x^2 - \frac{2x^3}{3})\Big|_0^2 = 1$$

$$\Rightarrow C(8 - \frac{16}{3}) = 1$$

$$\Rightarrow C = \frac{3}{8}$$

$$P(X > 1) = \int_{1}^{2} f(x)dx = \int_{1}^{2} C(4x - 2x^{2})dx$$

$$= C(2x^{2} - \frac{2x^{3}}{3})\Big|_{1}^{2} = 1$$

$$= C\left\{(8 - \frac{16}{3}) - (2 - \frac{2}{3})\right\}$$

$$= C\left\{\frac{8}{3} - \frac{4}{3}\right\}$$

$$= \frac{3}{8} \times \frac{4}{3}$$

$$= \frac{1}{2}.$$

Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

- Make a Graph of the above pdf.
- Find the probability that the technician will spend less than 30% of his workweek serving customers.
- Find the probability that the technician will spend 20% to 70% of hisworkweek serving customers.

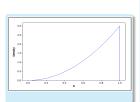
Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Make a Graph of the above pdf.

Find the probability that the technician will spend less than 30% of his workweek serving customers.

Find the probability that the technician will spend 20% to 70% of hisworkweek serving customers.



$$P(X < 0.3) = \int_0^{0.3} f(x) dx$$

$$= \int_1^2 3x^2 dx$$

$$= (x^3) \Big|_0^{0.3}$$

$$= (0.3)^3 - (0)^3$$

$$= 0.027$$

$$P(0.2 < X < 0.7) = \int_{0.2}^{0.7} f(x) dx$$

$$= \int_{0.2}^{0.7} 3x^2 dx$$

$$= (x^3) \Big|_{0.2}^{0.7}$$

$$= (0.7)^3 - (0.2)^3$$

$$= 0.337$$

### Exercise

Example: The amount of time in hours that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) := \begin{cases} 100 \ e^{-\frac{x}{100}} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

- What is the probability that a computer will function between 50 and 150 hours before breaking down?
- What is the probability that it will function for fewer than 100 hours?

### Exercise

Example: The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) := \begin{cases} \frac{100}{x^2} & \text{if } x > 100\\ 0 & \text{if } x \le 100. \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the fist 150 hours of operation? Assume

that the events  $E_i$ , i = 1, 2, 3, 4, 5, that the ith such tube will have to be replaced within this time are independent.

Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

 $f(x) := \begin{cases} 3x^2 & \text{if } 0 \le x \le 1 \\ 0 & \text{therwise.} \end{cases}$ 

- Obtain, F(x), the CDF of X.
- Use F(x) to compute P(0.5 < X < 0.8).

Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

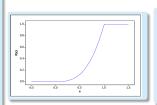
$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Obtain, F(x), the CDF of X and Graph it.

Use F(x) to compute P(0.5 < X < 0.8).

$$F(x) = P(X \le x) = \int_0^x f(y) dy$$
$$= \int_0^x 3y^2 dy$$
$$= (y^3) \Big|_0^x$$
$$= x^3$$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x^2 & \text{if } 0 \ge x < 1 \\ 1 & \text{if } x \ge 1. \end{cases}$$



$$P(0.5 < X < 0.8)$$
=  $F(0.8) - F(0.5)$   
=  $(0.8)^3 - (0.5)^3$   
=  $0.387$ 

- Example: Let X be a continuous random variable with Cumulative Distribution Function F(x), and density function f(x).
  - Obtain the cumulative distribution function of Y = 2X.
  - Obtain the probability density function of Y = 2X.

#### Outline

Percentiles, Quantiles, and Median

### Percentiles, Quantiles, and Median

#### Definition (Percentiles)

Let p be a number between 0 and 1. The  $(100)^{th}$  percentile of the distribution of a continuous random variable X, we shall denote by c, is that value for which

$$F\left( c \right) = p$$

i.e.  $\mathbf{c} = F^{-1}(p)$  . where  $F^{-1}(\cdot)$  is the inverse cumulative distribution function.

# Three Special Percentiles

Median The median of a continuous distribution, denoted by m, is the 50th percentile. So m satisfies

$$m = F^{-1}(0.5)$$

First Quartile The first quartile is defined to be

$$Q_1 = F^{-1}(0.25)$$

Third Quartile The third quartile is defined to be

$$Q_3 = F^{-1}(0.75)$$

Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

- find the median, and
- the interquartile range of the distribution.

Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

find the median, and

the interguartile range of the distribution.

We have already Shown

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x^2 & \text{if } 0 \ge x < 1 \\ 1 & \text{if } x \ge 1. \end{cases}$$

Note that if 
$$F(x) = y \implies x^3 = y \implies x = y^{\frac{1}{3}} \implies F^{-1}(y) = y^{\frac{1}{3}}$$
.

$$m = F^{-1}(0.5) = (0.5)^{\frac{1}{3}} = 0.794$$

IQR  
= 
$$Q_3 - Q_1$$
  
=  $F^{-1}(0.75) - F^{-1}(0.25)$   
=  $(0.75)^{\frac{1}{3}} - (0.25)^{\frac{1}{3}}$   
=  $0.909 - 0.630$   
=  $0.279$ 

#### Outline

Expected Value, Variance, & MGF of a Continuous Random Variable

#### Expected Value, or **mean** of a Continuous Random Variable

#### Definition (Expected Value or **mean** of a Continupues Random Variable)

If X is a continuous random variable with pdf f(x) on the support  $S_x$ , then the expected value (the mean) of X denoted by E(X) is given by

$$E(X) = \int_{\mathbb{S}_X} x \, f(x) \, dx \, ,$$

assuming the above integral exists.

E(X) is sometimes also denoted by  $\mu_{Y}$ 

#### Definition (Expected Value of a function of a Continuous Random Variable)

Let h(x) be any\* function. If X is a continuous random variable with pdf f(x) on the support  $\mathbb{S}_x$ , then the expected value h(X) denoted by E(h(X)) is given by

$$E(h(X)) = \int_{\mathbb{S}_X} h(x)f(x)dx,$$

assuming the above integral exists.

### Variance of a Random Variable

#### Definition (Variance a Random Variable)

Variance of a random variable X is defined to be

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

### A Few Properties of Expected Value and Variance of a Random Variable

#### Let a and b are constants, then

$$\bigcirc$$
 SD( $aX + b$ ) =  $|a|$ Var( $X$ )

# Moment Generating Function (mgf)

#### **Definition (Moment Generating Function)**

The Moment Generating Function (mgf) of X, denoted by  $\mathbf{M}_{x}(t)$  is deifined as

$$M_{\chi}(t) := E\left(e^{tX}\right)$$

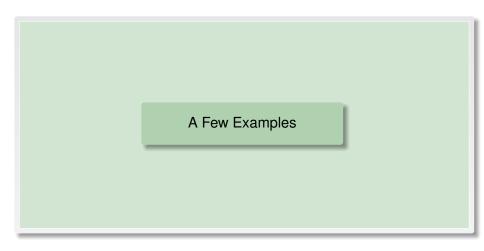
whenever it exists.

If a continuous random variable X has the probability density funciton  $f_X(x)$  for all  $x \in \mathbb{S}_X$ , the support of X, then assuming it exists

$$M_X(t) := E\left(e^{tX}\right) = \int e^{tx} f_X(x) dx.$$

$$\left\{x \in S_X\right\}$$

#### Outline



Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

- Find the expected value of percentage of time the technician spends serving customers.
- variance of percentage of time the technician spends serving customers.

Example: For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1\\ 0 & \text{otherwise.} \end{cases}$$

Find the expected value of percentage of time the technician spends serving customers. variance of percentage of time the technician spends serving customers.

$$E(X) = \int_{S_X} x f(x) dx$$

$$= \int_0^1 x (3x^2) dx$$

$$= \int_0^1 (3x^3) dx$$

$$= \frac{3x^4}{4} \Big|_0^1$$

$$= \frac{3}{4}$$

$$E(X^{2}) = \int_{\mathbb{S}_{X}} x^{2} f(x) dx$$

$$= \int_{0}^{1} x^{2} (3x^{2}) dx$$

$$= \int_{0}^{1} (3x^{4}) dx$$

$$= \frac{3x^{5}}{5} \Big|_{0}^{1}$$

$$= \frac{3}{5}$$

$$Var(X)$$
=  $E(X^2) - (E(X))^2$ 
=  $\frac{3}{5} - (\frac{3}{4})^2$ 
=  $0.6 - (0.75)^2$ 
=  $0.0375$ 

#### Example:

is

Find E(X) and Var(X) when the density function of X

$$f(x) := \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Example:

Find E(X) and Var(X) when the density function of X is

$$f(x) := \begin{cases} 2x & \text{if } 0 \le x \le 1\\ 0 & \text{otherwise.} \end{cases}$$

$$E(X) = \int_{\mathbb{S}_X} x f(x) dx$$

$$= \int_0^1 x(2x) dx$$

$$= \int_0^1 (2x^2) dx$$

$$= \frac{2x^3}{3} \Big|_0^1$$

$$= \frac{2}{3}$$

$$E(X^2) = \int_{\mathbb{S}_X} x^2 f(x) dx$$

$$= \int_0^1 x^2 (2x) dx$$

$$= \int_0^1 (2x^3) dx$$

$$= \frac{2x^4}{4} \Big|_0^1$$

$$= \frac{2}{4}$$

$$= \frac{1}{2}$$

$$\begin{aligned}
& \text{Var}(X) \\
&= E(X^2) - (E(X))^2 \\
&= \frac{1}{2} - \left(\frac{2}{3}\right)^2 \\
&= \frac{1}{2} - \frac{4}{9} \\
&= \frac{10}{9}
\end{aligned}$$

Example: Find  $E(e^X)$  and the Moment Generating Function for the continuous random variable with probability density function

$$f(x) := \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Example: Find  $E(e^X)$  and the Moment Generating Function for the continuous random variable with probability density function

$$f(x) := \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$E\left(e^{X}\right) = \int_{\mathbb{S}_{X}} e^{x} f(x) dx$$
$$= \int_{0}^{1} e^{x} (1) dx$$
$$= \left. e^{x} \right|_{0}^{1}$$
$$= \left. e^{1} - e^{0} \right.$$

$$M_X(t) := E\left(e^{tX}\right) = \int\limits_{S_X} e^{tx} f(x) dx$$

$$= \int_0^1 e^{tx} (1) dx$$

$$= \frac{1}{t} e^{tx} \Big|_0^1$$

$$= \frac{1}{t} e^t - \frac{1}{t} e^0$$

$$= \frac{e^t - 1}{t}$$

#### Example:

Let X denote the resistance of a randomly chosen resistor, and suppose that its pdf is given by

$$f(x) := \begin{cases} \frac{x}{18} & \text{if } 8 \le x \le 10\\ 0 & \text{otherwise.} \end{cases}$$



Find and graph the cdf of X.

Find  $P(8.6 < X \le 9.8)$ .

Find the median of the resistance of such resistors. Find the mean and variance of X.

Example: The length of time to failure (in hundreds of hours) for a transistor is a random variable X with cumulative distribution function given by

$$F(x) := \begin{cases} 1 - e^{-x^2} & \text{for } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Find a pdf of X f(x).

Find the probability that the transistor operates for at least 200 hours.

Find the 30<sup>th</sup> percentile of X.

#### Example: given by

Weekly CPU time used by an accounting firm has probability density function (measured in hours)

$$f(x) := \begin{cases} \frac{3}{64}x^2(4-X) & \text{for } 0 \le x \le 4\\ 0 & \text{otherwise.} \end{cases}$$



Find the F(x) for weekly CPU time.

Find the probability that the of weekly CPU time will exceed two hours for a selected week.

Find the expected value and variance of weekly CPU time.

Find the probability that the of weekly CPU time will be within half an hour of the expected weekly CPU time.

The CPU time costs the firm \$200 per hour. Find the expected value and variance of the weekly cost for CPU time.

#### Example:

The length of time required by students to complete a one-hour exam is a random variable with a density function given by

$$f(x) := \begin{cases} cy^2 + y & \text{for } 0 \le y \le 1\\ 0 & \text{otherwise.} \end{cases}$$

Find the F(v)

Find c that makes this function a valid probability density function.

Find the probability that a randomly selected student will finish in less than half an hour.

Find the time that 95% of the students finish before it.

Given that a particular student needs at least 15 minutes to complete the exam, find the probability that she will require at least 30 minutes to finish.

