STAT 320: Principles of Probability

Unit 6 Part:A

Continuous Random Variables

United Arab Emirates University

Department of Statistics

Outline

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

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

CDF: Example

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

x	0	1	2	3
$p_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

The cdf of X is:

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

Example of CDF

x	0	1	2	3
$P_X(X=x)$	$\frac{1}{8}=.125$	$\frac{3}{8}=.375$	$\frac{3}{8}=0.375$	$\frac{1}{8}=.125$

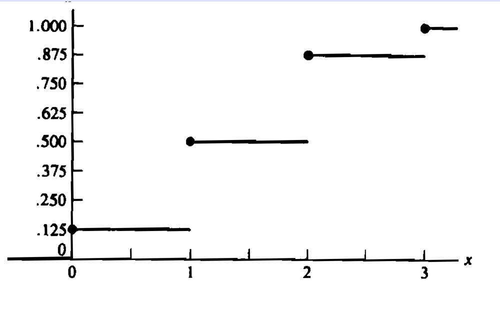


Figure: The polt of $F_X(x)$: CDF of the random variable X

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, $2.5 \notin \mathbb{S}X$, however

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

Characterization of *any* CDF Function

Characterization of a CDF

Theorem

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

1. $\lim_{x \to -\infty} F(x) = 0$ *and* $\lim_{x \to \infty} F(x) = 1$.
2. *F(x) is a nondecreasing function of x*
3. *F(x) is right-continuous; that is, for every real number x_0,* $\lim_{x \downarrow x_0} F(x) = F(x_0)$.

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

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

Example: CDF continuous

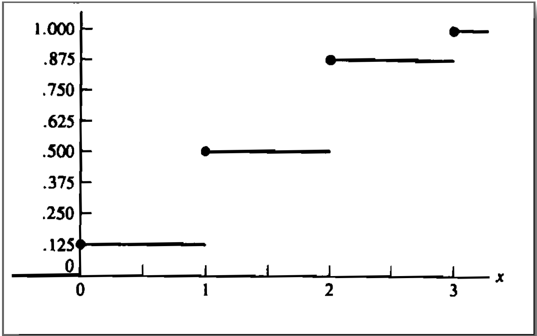


Figure: The polt of $F_X(x)$: CDF of the random variable X

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

$P(X = 2) = \lim_{x \downarrow 2} F_X(X) - \lim_{x \uparrow 2} F_X(x) = 0.5 - 0.125 = 0.375.$

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

Example: CDF continuous

Example: An example of a continuous cdf is the function

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

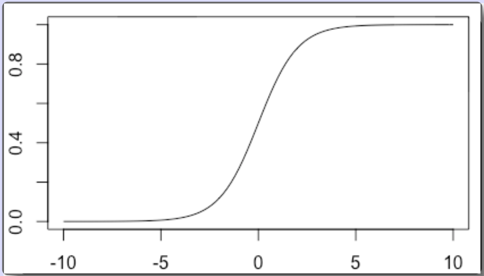


Figure: The polt of $F_X(x)$: CDF of the random variable X

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

Example:

Question : : Prove that the following functions are valid cdfs.

1. $F(x) = e^{-e^{-x}}$ for all $x \in \mathbb{R}$.
2. $F(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x)$ for all $x \in \mathbb{R}$.

Definition (Discrete Random Variable)

A random variable X is discrete if it's support $\mathbb{S}_x$ 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 cdf $F_x(x)$ is a continuous function of x.

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 finite support?

Example: CDF continuous

Example: An example of a continuous cdf is the function

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

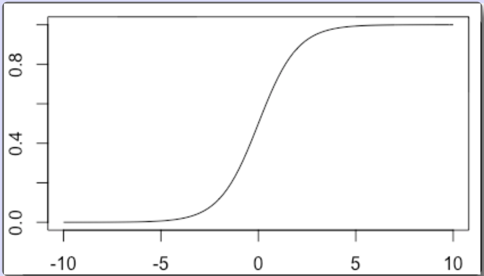


Figure: The polt of $F_X(x)$: CDF of the random variable X

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

Example: CDF continuous

Example of CDF of a Continuous Random Variable:

$$F_X(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ 1 - e^{-x} & \text{if } x > 0 \end{cases}$$

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

Example:

Question : : Prove that the following functions are valid cdfs.

1. $F(x) = e^{-e^{-x}}$ for all $x \in \mathbb{R}$.
2. $F(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x)$ for all $x \in \mathbb{R}$.

Probability Density Function (pmf): For continuous RV

Definition (Probability Density Function)

The probability density function or pdf, $f_x(x)$, of a continuous random variable X is the function that satisfies

$$F_X(x) = \int_{-\infty}^{x} f_x(x)dx$$

Comment: Using the Fundamental Theorem of Calculus, if $f_X(x)$ is continuous, we have the further relationship $f_X(x) = \dfrac{d}{dx}F_X(x)$.

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

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x)dx.$$

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

1. $f(x) \geq 0$ for all x.
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a \leq X \leq b) = \int_a^b f(x)dx$ for all $a < b$.

If X is a continuous random variable then,

- $P(X = c) = 0$ for any $c \in \mathbb{R}$
- $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$.

Example

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

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

a). What is the value of C?

b). Find $P(X > 1)$.

Example

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

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

a). What is the value of C?

b). Find $P(X > 1)$.

According to the property of the pdf

$$\begin{aligned} & \int f(x)dx = 1 \\ \Longrightarrow \quad & \int_0^2 C(4x - 2x^2)dx = 1 \\ \Longrightarrow \quad & C(2x^2 - \frac{2x^3}{3})\Big|_0^2 = 1 \\ \Longrightarrow \quad & C(8 - \frac{16}{3}) = 1 \\ \Longrightarrow \quad & C = \frac{3}{8} \end{aligned}$$

$$\begin{aligned} P(X > 1) &= \int_1^2 f(x)dx \\ &= \int_1^2 C(4x - 2x^2)dx \\ &= C(2x^2 - \frac{2x^3}{3})\Big|_1^2 = 1 \\ &= C\left\{(8 - \frac{16}{3}) - (2 - \frac{2}{3})\right\} \\ &= C\left\{\frac{8}{3} - \frac{4}{3}\right\} \\ &= \frac{3}{8} \times \frac{4}{3} \\ &= \frac{1}{2}. \end{aligned}$$

Example

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

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

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

Example

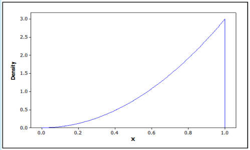
Example : For a given IT technician in a support center, let X denote the percentage of time, out of a 40-hour work week, that he is directly serving customers. Suppose that X has a probability density function given by

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

a). Make a Graph of the above pdf.

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

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



$$\begin{aligned} P(X < 0.3) &= \int_0^{0.3} f(x)dx \\ &= \int_1^{2} 3x^2 dx \\ &= (x^3)\Big|_0^{0.3} \\ &= (0.3)^3 - (0)^3 \\ &= 0.027 \end{aligned}$$

$$\begin{aligned} P(0.2 < X < 0.7) &= \int_{0.2}^{0.7} f(x)dx \\ &= \int_{0.2}^{0.7} 3x^2 dx \\ &= (x^3)\Big|_{0.2}^{0.7} \\ &= (0.7)^3 - (0.2)^3 \\ &= 0.337 \end{aligned}$$

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

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

What is the probability that

a). a computer will function between 50 and 150 hours before breaking down?

b). 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

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

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.

Cumulative Distribution Function (cdf)

Definition ((The Cumulative Distribution Function))

The cumulative distribution function (cdf) F(x) of a continuous random variable X with pdf $f()$ is defined for every number x by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx.$$

A Few Properties of F(x)

1. $P(a < X \leq b) = F(b) - F(a)$.
2. $P(X > b) = 1 - F(b)$
3. If X is a continuous random variable with cdf $F(x)$ then at every x at which $\frac{dF(x)}{dx}$ exists: $f(x) = \frac{dF(x)}{dx}$

Example

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

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

a). Obtain, $F(x)$, the CDF of X.

b). Use $F(x)$ to compute $P(0.5 < X \le 0.8)$.

:

Example

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

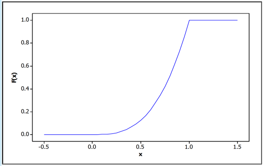
$$f(x) := \begin{cases} 3x^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

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

b). Use $F(x)$ to compute $P(0.5 < X \leq 0.8)$.

$$\begin{aligned} F(x) = P(X \leq x) &= \int_0^x f(y)dy \\ &= \int_0^x 3y^2 dy \\ &= (y^3)\Big|_0^x \\ &= x^3 \end{aligned}$$

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



$$\begin{aligned} & P(0.5 < X < 0.8) \\ = \;& F(0.8) - F(0.5) \\ = \;& (0.8)^3 - (0.5)^3 \\ = \;& 0.387 \end{aligned}$$

Example

Example : Let X be a continuous random variable with Cumulative Distribution Function F(x), and density function f(x).

1. Obtain the cumulative distribution function of $Y = 2X$.
2. Obtain the probability density function of $Y = 2X$.

:

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

$$F(c) = p$$

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

Special percentiles:

1. The median of a continuous distribution, denoted by m, is the 50th percentile. So m satisfies $m = F^{-1}(0.5)$.
2. The first and the third quartiles can be computed as $Q_1 = F^{-1}(0.25)$
3. The third and the third quartiles can be computed as $Q_3 = F^{-1}(0.75)$

Example

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

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

a). find the median, and

b). the interquartile range of the distribution.

Example

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

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

a). find the median, and

b). the interquartile range of the distribution.

We have already Shown

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

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

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

$$\begin{aligned} & \text{IQR} \\ = \; & Q_3 - Q_1 \\ = \; & F^{-1}(0.75) - F^{-1}(0.25) \\ = \; & (0.75)^{\frac{1}{3}} - (0.25)^{\frac{1}{3}} \\ = \; & 0.909 - 0.630 \\ = \; & 0.279 \end{aligned}$$

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 $f(x)$, then the expected value (the mean) of X denoted by $E(X)$ or μ_X is given by

$$\mu_X := E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

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 $f(x)$, then the expected value $h(X)$ denoted by $E(h(X))$ is given by

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

Variance of a Random Variable

Definition (Variance a Random Variable)

Variance of a random variable X is defined to be

$$\text{Var}(X) = E(X^2) - (E(X))^2$$

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

Let a and b are constants, then

1. $E(a + bX) = a + bE(X)$
2. $\mathrm{Var}(a + bX) = b^2\mathrm{Var}(X)$

Outline

A Few Examples

Example

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

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

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

b). variance of percentage of time the technician spends serving customers.

Example

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

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

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

b). variance of percentage of time the technician spends serving customers.

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_0^1 x(3x^2)dx \\ &= \int_0^1 (3x^3)dx \\ &= \frac{3x^4}{4}\Big|_0^1 \\ &= \frac{3}{4} \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx \\ &= \int_0^1 x^2(3x^2)dx \\ &= \int_0^1 (3x^4)dx \\ &= \frac{3x^5}{5}\Big|_0^1 \\ &= \frac{3}{5} \end{aligned}$$

$$\begin{aligned} &\mathrm{Var}(X) \\ &= E(X^2) - (E(X))^2 \\ &= \frac{3}{5} - \left(\frac{3}{4}\right)^2 \\ &= 0.6 - (0.75)^2 \\ &= 0.0375 \end{aligned}$$

Example

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

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

Example

Example : Find $E(X)$ and $\mathrm{Var}(X)$ when the density function of X is

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

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_0^1 x(2x) dx \\ &= \int_0^1 (2x^2) dx \\ &= \frac{2x^3}{3}\Big|_0^1 \\ &= \frac{2}{3} \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx \\ &= \int_0^1 x^2 (2x) dx \\ &= \int_0^1 (2x^3) dx \\ &= \frac{2x^4}{4}\Big|_0^1 \\ &= \frac{2}{4} \\ &= \frac{1}{2} \end{aligned}$$

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

Example

Example : Find $E(e^X)$ when the density function of X is

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

Example

Example : Find $E(e^X)$ when the density function of X is

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

$$\begin{aligned} E(e^X) &= \int_{-\infty}^{\infty} e^x f(x)dx \\ &= \int_0^1 e^x (1)dx \\ &= e^x \Big|_0^1 \\ &= e^1 - e^0 \\ &= e - 1 \end{aligned}$$

Exercise

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

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

1. Find and graph the cdf of X.
2. Find $P(8.6 < X \leq 9:8)$.
3. Find the median of the resistance of such resistors.
4. Find the mean and variance of X.

Exercise

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

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

1. Find a pdf of X $f(x)$.
2. Find the probability that the transistor operates for at least 200 hours.
3. Find the 30^{th} percentile of X.

Exercise

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

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

1. Find the F(x) for weekly CPU time.
2. Find the probability that the of weekly CPU time will exceed two hours for a selected week.
3. Find the expected value and variance of weekly CPU time.
4. Find the probability that the of weekly CPU time will be within half an hour of the expected weekly CPU time.
5. The CPU time costs the firm $200 per hour. Find the expected value and variance of the weekly cost for CPU time.

Exercise

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

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

1. Find c that makes this function a valid probability density function.
2. Find the F(y)
3. Graph f(y) and F(y).
4. Find the probability that a randomly selected student will finish in less than half an hour.
5. Find the time that 95% of the students finish before it.
6. 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.

Questions?