CONTINUOUS RANDOM VARIABLES

MSc. Nguyễn Hoàng Huy Tú

November 24, 2023

Remark

Remark

Let X be a continuous random variable. Then

- X can take any value in an interval $\Rightarrow$ infinite possible values.
- $P(X = x) = 0$. That means the probability of a specific value is 0 for continuous random variables.
- $P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X < b)$

Cumulative Distribution Function

Property

The cumulative distribution function, $F(x)$, for a continuous random variable X expresses the probability that X does not exceed the value of x, as a function of x:

$$F(x) = P(X \leq x)$$

Then

- $P(a < X < b) = F(b) - F(a)$
- $F(-\infty) = 0$
- $F(+\infty) = 1$

Example

Consider a gasoline station that has a 1,000-gallon storage tank that is filled each morning at the start of the business day. The random variable X indicates the gasoline sales in gallons for a particular day. Given the cumulative distribution as follows

$$F(x) = \begin{cases} 0; & x < 0 \\ 0,001x; & 0 \leq x \leq 1000 \\ 1; & x > 1000 \end{cases}$$

a. Find the probability of sales is between 250 and 750.

b. Find the probabitlity that the sales is greater than or equal to 400?

Probability Density Function

Definition

Let X be a continuous random variable. Suppose $F(x)$ is the cumulative distribution function. The the derivative of $F(x)$ is called a probability density function, denoted by $f(x)$. That is,

$$f(x) = F'(x)$$

For example, given the cumulative distribution function

$$F(x) = \begin{cases} 0; & x < 0 \\ 0,001x; & 0 \leq x \leq 1000 \\ 1; & x > 1000 \end{cases}$$

Then the probability density function is

$$f(x) = \begin{cases} 0,001; & 0 \leq x \leq 1000 \\ 0; & x \notin [0, 1] \end{cases}$$

Probability Density Function

Property

Let $f(x)$ be the probability density function of a continuous random variable. Then

- $f(x) > 0$ for all values of x.
- $P(a < X < b) = \int_a^b f(x)dx$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$. That implies the total area under the curve $f(x)$ is 1.

Example

A repair team is responsible for a stretch of oil pipeline 2 miles long. The distance (in miles) at which any fracture occurs can be represented by a uniformly distributed random variable, with probability density function

$$f(x) = 0.5$$

Find the probability that any given fracture occurs between 0.5 mile and 1.5 miles along this stretch of pipeline.

Exercise

Ex: The jurisdiction of a rescue team includes emergencies occurring on a stretch of river that is 4 miles long. If X denotes the distance (in miles) of an emergency from the northernmost point of this stretch of river, its probability density function is as follows:

$$f(x) = \begin{cases} 0,25; & 0 < x < 4 \\ 0; & x \notin (0,4) \end{cases}$$

a. Find the probability that a given emergency arises within 1 mile of the northernmost point of this stretch of river.
b. The rescue team's base is at the midpoint of this stretch of river. Find the probability that a given emergency arises more than 1.5 miles from this base.

The Uniform Distribution

Definition

The continuous random variable X is said to follow the Uniform distribution if the probability density function is as follows

$$f(x) = \begin{cases} \dfrac{1}{b-a}; & a \leq x \leq b \\ 0; & x \notin [a, b] \end{cases}$$

Expectation for Continuous Random Variables

Formula

Let X be a continuous random variable and $f(x)$ its probability density function. Then

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

and

$$\sigma_X^2 = Var(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

The standard deviation of X, σ_X, is the square root of the variance.

Example

The random variable X has probability density function as follows:

$$f(x) = \begin{cases} x; & 0 < x < 1 \\ 2 - x & 1 < x < 2 \\ 0 & \text{for all other values of x} \end{cases}$$

a. Find the probability that X takes a value between 0,2 and 1,2.

b. Find the mean and the variance of X.

Expectation for Continuous Random Variables

Property

Let X be a continuous random variable with mean μ_X and variance σ_X^2 and let a and b be any constant fixed numbers. Define the random variable W as follows:

$$W = a + bX$$

Then

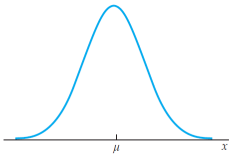
$$\mu_W = E(a + bX) = a + b\mu_X$$

and

$$\sigma_W^2 = Var(a + bX) = b^2\sigma_X^2$$

Normal Distribution

There is a wide range of continuous random variables which follow the normal distribution such as the weight of food packages, total sales or production, stock prices, …Their probability density functions have the same pattern as follows



The shape of density function

- is bell-shaped;
- is symmetric;
- has a central tendency (indicated by the mean μ)

Normal Distribution

Definition

A continuous random variable X is said to follow the normal distribution if its probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\dfrac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$$

A normal distribution only depends on two parameters

- $\mu = E(X)$: The mean or the expected value;
- $\sigma^2 = Var(X)$: The variance.

We denote

$$X \sim N(\mu, \sigma^2)$$

The Standard Normal Distribution

Definition

A random variable Z is said to follow the standard normal distribution if it follows the normal distribution with the mean 0 and variance 1. That is,

$$Z \sim N(0, 1)$$

Remark: Let $X \sim N(\mu, \sigma^2)$.

Set $Z = \dfrac{X - \mu}{\sigma}$. Then

$$Z \sim N(0, 1).$$

Cumulative Distribution

Let $Z \sim N(0,1)$. Then its density function is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Since $f(x) = F'(x)$, we have $F(x) = \int_{-\infty}^{x} f(x)dx$.

Definition

The cumulative distribution function of a random variable, which follows the standard normal distribution, is

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

The Value of Cumulative Distribution Function

The value of the cumulative distribution function

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

can be calculated by computer or found in Appendix Table 1.

- For $0 \leq x < 4$, $F(x)$ can be found in Appendix Table 1.
- For $-4 < x < 0$, $F(x) = 1 - F(-x)$.
- For $x \geq 4$, $F(x) \approx 1$.
- For $x \leq -4$, $F(x) \approx 0$.

Example

Let $Z \sim N(0,1)$. Given the cumulative distribution function of Z, that is $F(x)$. Find

a) $F(-\frac{1}{3}) =?$
b) $F(1,2) =?$
c) $F(2,33) =?$
d) $F(-1,7) =?$
e) $F(-5) =?$

Formula

Let $X \sim N(\mu, \sigma^2)$. We have

$$\begin{aligned} P(a < X < b) &= P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) \\ &= P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right) \\ &= F\left(\frac{b-\mu}{\sigma}\right) - F\left(\frac{a-\mu}{\sigma}\right) \end{aligned}$$

Formula

Let X follow the normal distribution with the mean μ and variance σ^2. Then

$$P(a < X < b) = F\left(\frac{b-\mu}{\sigma}\right) - F\left(\frac{a-\mu}{\sigma}\right)$$

and the value of $F(x)$ can be found in Appendix Table 1.

Example

Example: Southwest Co-op produces bags of fertilizer, and it is concerned about impurity content. It is believed that the weights of impurities per bag are normally distributed with a mean of 12.2 grams and a standard deviation of 2.8 grams. A bag is chosen at random.

a. What is the probability that it contains between 12 and 15 grams of impurities?

b. What is the probability that it contains less than 10 grams of impurities?

c. What is the probability that it contains more than 15 grams of impurities?

Example

Example: A furniture manufacturer has found that the time spent by workers assembling a particular table follows a normal distribution with a mean of 150 minutes and a standard deviation of 40 minutes.

a. The probability is 0.9 that a randomly chosen table requires more than how many minutes to assemble?

b. The probability is 0.8 that a randomly chosen table can be assembled in fewer than how many minutes?

c. Five tables are chosen at random. What is the probability that at least one of them requires at least 2 hours to assemble?

Exercise

Ex 1: The company has developed a range feeding program with organic grain supplements to produce their product. The mean weight of its frozen turkeys is 15 pounds with a variance of 4. Historical experience indicates that weights can be approximated by the normal probability distribution. What percentage of the company's turkey units will be over 18 pounds?

Ex 2: The company's sales experience indicates that daily cell phone sales in its stores follow a normal distribution with a mean of 60 and a standard deviation of 15. The marketing department conducts a number of routine analyses of sales data to monitor sales performance.
a. What proportion of store sales days will have sales between 85 and 95 given that sales are following the historical experience?
b. Find the cutoff point for the top 10% of all daily sales.

Exercise

Ex 3: The lifespan of a product is a normally distributed random variable with the mean 4.2 years and the standard deviation 1.5 years. If you sell a product, you make a profit of 100$, but if the product has a warranty, you lose 300$. So for the average profit when selling each product to be 30$, what is the warranty period?

Linear combination of normally distributed random variables

Theorem

Let $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$. Suppose that X, Y are independent and

$$W = aX + bY.$$

Then W also follows the normal distribution. In particular,

$$W \sim N\left(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2\right).$$

Normal Distribution Approximation for Binomial Distribution

SELF-STUDY