

Lecture 5. Continuous Probability Distribution

Vu Nguyen Son Tung

Faculty of information technology
HANU

Continuous Random Variable

A random variable is called a **continuous random variable** if

- If a sample space contains an infinite number of possibilities equal to the number of points on a line segment.
- $P\{X = a\} = 0, \quad \forall a \in \mathbb{R}.$

Examples. Continuous random variables represent measured data, such as all possible heights, weights, temperatures, distance, or life periods, whereas discrete random variables represent count data, such as the number of defectives in a sample of k items or the number of highway fatalities per year in a given state.

Probability Density Function

Although the probability distribution of a continuous random variable cannot be presented in tabular form, it can be stated as a formula. Such a formula would necessarily be a function of the numerical values of the continuous random variable X and as such will be represented by the functional notation $f(x)$. In dealing with continuous variables, $f(x)$ is usually called the **probability density function**, or simply **the density function**, of X .

The function $f(x)$ is a probability density function (pdf) for the continuous random variable X , defined over the set of real numbers, if

- i) $f(x) \geq 0, \forall x \in \mathbb{R};$
- ii) $\int_{-\infty}^{+\infty} f(x)dx = 1;$
- iii) $\forall a < b: P\{a < X < b\} = \int_a^b f(x)dx.$

Note: Since $P\{X = a\} = P\{X = b\} = 0$, so
 $P\{a \leq X \leq b\} = P\{a \leq X < b\} = P\{a < X < b\} = P\{a < X \leq b\}.$

Example 1. Given the continuous X , which has $f(x) = \begin{cases} 0, & x < 1; \\ \frac{1}{x^2}, & x > 1. \end{cases}$

Compute $P\{2 < X < 3\}$.

Example 1. Given the continuous X , which has $f(x) = \begin{cases} 0, & x < 1; \\ \frac{1}{x^2}, & x > 1. \end{cases}$

Compute $P\{2 < X < 3\}$.

Solution.

$$P\{2 < X < 3\} = \int_2^3 f(x) dx = \int_2^3 \frac{1}{x^2} dx = \frac{1}{6}.$$

Example 1. Given the continuous X , which has $f(x) = \begin{cases} 0, & x < 1; \\ \frac{1}{x^2}, & x > 1. \end{cases}$

Compute $P\{2 < X < 3\}$.

Solution.

$$P\{2 < X < 3\} = \int_2^3 f(x) dx = \int_2^3 \frac{1}{x^2} dx = \frac{1}{6}.$$

Example 2. Given the continuous X : $f(x) = \begin{cases} 1 + x, & -1 \leq x < 0; \\ 1 - x, & 0 < x \leq 1; \\ 0, & |x| > 1. \end{cases}$

Compute $P\{-1/2 < X < 1\}$.

Example 1. Given the continuous X , which has $f(x) = \begin{cases} 0, & x < 1; \\ \frac{1}{x^2}, & x > 1. \end{cases}$

Compute $P\{2 < X < 3\}$.

Solution.

$$P\{2 < X < 3\} = \int_2^3 f(x) dx = \int_2^3 \frac{1}{x^2} dx = \frac{1}{6}.$$

Example 2. Given the continuous X : $f(x) = \begin{cases} 1 + x, & -1 \leq x < 0; \\ 1 - x, & 0 < x \leq 1; \\ 0, & |x| > 1. \end{cases}$

Compute $P\{-1/2 < X < 1\}$.

Solution.

$$P\{-1/2 < X < 1\} = \int_{-\frac{1}{2}}^1 f(x) dx = \int_{-\frac{1}{2}}^0 (1 + x) dx + \int_0^1 (1 - x) dx = \frac{7}{8}.$$

Cumulative Distribution Function

Definition. The **cumulative distribution function** $F(x)$ of a continuous random variable X with the density function $f(x)$ is

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

Properties.

- $0 \leq F(x) \leq 1$;
- $\forall x_1 \leq x_2 \implies F(x_1) \leq F(x_2)$;
- $F(x)$ is continuous;
- $\lim_{x \rightarrow +\infty} F(x) = P\{X < +\infty\} = 1$ and $\lim_{x \rightarrow -\infty} F(x) = P\{X < -\infty\} = 0$;
- $f(x) = F'(x)$ and $F(x) = \int_{-\infty}^x f(t)dt$.

Example 3. Given the continuous X , such that $f(x) = \begin{cases} 0, & x < 0; \\ \frac{6x}{5}, & 0 \leq x \leq 1; \\ \frac{6}{5x^4}, & x > 1. \end{cases}$

Find the function $F(x)$ of X .

Example 3. Given the continuous X , such that $f(x) = \begin{cases} 0, & x < 0; \\ \frac{6x}{5}, & 0 \leq x \leq 1; \\ \frac{6}{5x^4}, & x > 1. \end{cases}$

Find the function $F(x)$ of X .

Solution.

- If $x < 0$: $F(x) = \int_{-\infty}^x f(s) ds = 0.$
- If $0 \leq x \leq 1$: $F(x) = \int_{-\infty}^x f(s) ds = \int_{-\infty}^0 0 ds + \int_0^x \frac{6s}{5} ds = \frac{3}{5}x^2.$
- If $x > 1$: $F(x) = \int_{-\infty}^x f(s) ds = \int_0^1 \frac{6s}{5} ds + \int_1^x \frac{6}{5s^4} ds = 1 - \frac{2}{5x^3}.$

Therefore

$$F(x) = \begin{cases} 0, & x < 0; \\ \frac{3x^2}{5}, & 0 \leq x \leq 1; \\ 1 - \frac{2}{5x^3}, & x > 1. \end{cases}$$

Example 4. The time (minutes) for a person waiting in line to pay at the supermarket is a continuous random variable X with a function:

$$F(x) = \begin{cases} 0, & x \leq 0; \\ ax^2, & 0 < x < 3; \\ 1, & x \geq 3. \end{cases}$$

- a) Find a and probability density function X .
- b) Calculate the probability that out of 3 people in line, 2 people have to wait no more than 2 minutes.

Example 4. The time (minutes) for a person waiting in line to pay at the supermarket is a continuous random variable X with a function:

$$F(x) = \begin{cases} 0, & x \leq 0; \\ ax^2, & 0 < x < 3; \\ 1, & x \geq 3. \end{cases}$$

- a) Find a and probability density function X .
b) Calculate the probability that out of 3 people in line, 2 people have to wait no more than 2 minutes.

Solution.

- Since $\lim_{x \rightarrow 3^-} F(x) = F(3) \implies \lim_{x \rightarrow 3^-} ax^2 = 1$. therefore $A = 1/9$. And we have

$$f(x) = F'(x) = \begin{cases} 0, & x \notin (0, 3] \\ \frac{2}{9}x, & 0 < x < 3. \end{cases}$$

Example 4. The time (minutes) for a person waiting in line to pay at the supermarket is a continuous random variable X with a function:

$$F(x) = \begin{cases} 0, & x \leq 0; \\ ax^2, & 0 < x < 3; \\ 1, & x \geq 3. \end{cases}$$

- a) Find a and probability density function X .
b) Calculate the probability that out of 3 people in line, 2 people have to wait no more than 2 minutes.

Solution.

- Since $\lim_{x \rightarrow 3^-} F(x) = F(3) \implies \lim_{x \rightarrow 3^-} ax^2 = 1$. therefore $A = 1/9$. And we have

$$f(x) = F'(x) = \begin{cases} 0, & x \notin (0, 3] \\ \frac{2}{9}x, & 0 < x < 3. \end{cases}$$

- The probability that a customer have to wait no more than 2 minutes is:

$$P(X \leq 2) = F(2) = \frac{4}{9}.$$

So the probability that out of 3 people, 2 people have to wait no more than 2 minutes is: $P_2(2) = C_3^2 \left(\frac{4}{9}\right)^2 \left(\frac{5}{9}\right)^1 = 0,329$.

CHARACTERISTICS OF RANDOM VARIABLE

Given the continuous X with density $f(x)$. Then

Mean of X , denote by EX , is a number determined by

$$EX = \int_{-\infty}^{+\infty} xf(x)dx$$

Variance of X , denoted by σ^2 , is a number determined by

$$DX = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x)dx = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mu^2, \quad \mu \equiv EX.$$

Standard deviation of X is $\sigma_X = \sqrt{\sigma^2}$.

Mode of X is the value x_0 , denoted by $\text{mod } X$, if x_0 is the maximum point of the density function $f(x)$.

Median of X is number m if $P\{X < m\} = P\{X > m\}$ or $F(m) = \frac{1}{2}$

Example 5. Given a continuous X with:

$$f(x) = \begin{cases} ax^3, & 0 \leq x \leq 3; \\ 0, & x \notin [0, 3]. \end{cases}$$

Find a , mean, variance, mode, median of X .

Example 5. Given a continuous X with:

$$f(x) = \begin{cases} ax^3, & 0 \leq x \leq 3; \\ 0, & x \notin [0, 3]. \end{cases}$$

Find a , mean, variance, mode, median of X .

Solution.

- Following the definition:

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \implies 1 = \int_0^3 ax^3 dx = \frac{81}{4} a \implies a = \frac{4}{81}$$

Example 5. Given a continuous X with:

$$f(x) = \begin{cases} ax^3, & 0 \leq x \leq 3; \\ 0, & x \notin [0, 3]. \end{cases}$$

Find a , mean, variance, mode, median of X .

Solution.

- Following the definition:

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \implies 1 = \int_0^3 ax^3 dx = \frac{81}{4} a \implies a = \frac{4}{81}$$

- $EX = \int_0^3 x \frac{4}{81} x^3 dx = 2, 4.$

Example 5. Given a continuous X with:

$$f(x) = \begin{cases} ax^3, & 0 \leq x \leq 3; \\ 0, & x \notin [0, 3]. \end{cases}$$

Find a , mean, variance, mode, median of X .

Solution.

- Following the definition:

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \implies 1 = \int_0^3 ax^3 dx = \frac{81}{4} a \implies a = \frac{4}{81}$$

- $EX = \int_0^3 x \frac{4}{81} x^3 dx = 2,4.$

- Ta có:

$$\int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^3 x^2 \frac{4}{81} x^3 dx = 6 \Rightarrow DX = 6 - (EX)^2 = 6 - (2,4)^2 = 0,24$$

Example 5. Given a continuous X with:

$$f(x) = \begin{cases} ax^3, & 0 \leq x \leq 3; \\ 0, & x \notin [0, 3]. \end{cases}$$

Find a , mean, variance, mode, median of X .

Solution.

- Following the definition:

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \implies 1 = \int_0^3 ax^3 dx = \frac{81}{4} a \implies a = \frac{4}{81}$$

- $EX = \int_0^3 x \frac{4}{81} x^3 dx = 2, 4.$

- Ta có:

$$\int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^3 x^2 \frac{4}{81} x^3 dx = 6 \Rightarrow DX = 6 - (EX)^2 = 6 - (2, 4)^2 = 0, 24$$

- It's easy to see that $f(x)$ has the maximum point $x = 3$ therefore mod $X = 3$.

Example 5. Given a continuous X with:

$$f(x) = \begin{cases} ax^3, & 0 \leq x \leq 3; \\ 0, & x \notin [0, 3]. \end{cases}$$

Find a , mean, variance, mode, median of X .

Solution.

- Following the definition:

$$\int_{-\infty}^{+\infty} f(x) dx = 1 \implies 1 = \int_0^3 ax^3 dx = \frac{81}{4} a \implies a = \frac{4}{81}$$

- $EX = \int_0^3 x \frac{4}{81} x^3 dx = 2, 4.$

- Ta có:

$$\int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^3 x^2 \frac{4}{81} x^3 dx = 6 \implies DX = 6 - (EX)^2 = 6 - (2, 4)^2 = 0, 24$$

- It's easy to see that $f(x)$ has the maximum point $x = 3$ therefore mod $X = 3$.
- Denote m by median of X . We have

$$F(m) = P\{X < m\} = \frac{1}{2} \implies \int_0^m \frac{4}{81} x^3 dx = \frac{1}{2} \implies \frac{m^4}{81} = \frac{1}{2} \implies m = 2, 523.$$

Random Variables Related To The Given Random Variable X

Consider a new random variable $g(X)$, which depends on X ; that is, each value of $g(X)$ is determined by the value of X .

This result is generalized in the below Theorem for both discrete and continuous random variables.

Random Variables Related To The Given Random Variable X

Consider a new random variable $g(X)$, which depends on X ; that is, each value of $g(X)$ is determined by the value of X .

This result is generalized in the below Theorem for both discrete and continuous random variables.

Let X be a random variable with probability distribution $f(x)$. The expected value and variance of the random variable $g(X)$ is

$$\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x), \quad \sigma_{g(X)}^2 = \sum_x [g(x) - \mu_{g(X)}]^2 f(x).$$

if X is discrete, and

$$\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx, \quad \sigma_{g(X)}^2 = \int_{-\infty}^{\infty} [g(x) - \mu_{g(X)}]^2 f(x)dx.$$

if X is continuous.

Example 6. Find the density of $Y = X^3$ by using the given $f(x)$ of X .

Solution. Since

$$F_Y(x) = P\{Y < x\} = P\{X^3 < x\} = P\{X < x^{1/3}\} = F(x^{1/3}) \quad \text{therefore} \\ f_Y(x) = F'_Y(x) = F'(x^{1/3}) \frac{1}{3} x^{-2/3} = 1/3 x^{-2/3} f(x^{1/3}).$$

Example 6. Find the density of $Y = X^3$ by using the given $f(x)$ of X .

Solution. Since

$$F_Y(x) = P\{Y < x\} = P\{X^3 < x\} = P\{X < x^{1/3}\} = F(x^{1/3}) \quad \text{therefore} \\ f_Y(x) = F'_Y(x) = F'(x^{1/3}) \frac{1}{3} x^{-2/3} = \frac{1}{3} x^{-2/3} f(x^{1/3}).$$

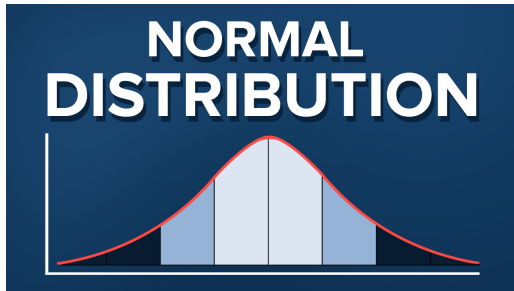
Example 7.

- Find the density function of $Y = X^3$ if $f(x)$ of X is given in the example 3.
- Compute $P\{10 < Y < 20\}$;
- Compute EY .

$$f(x) = \begin{cases} 0, & x < 0; \\ \frac{6x}{5}, & 0 \leq x \leq 1; \\ \frac{6}{5x^4}, & x > 1. \end{cases}$$

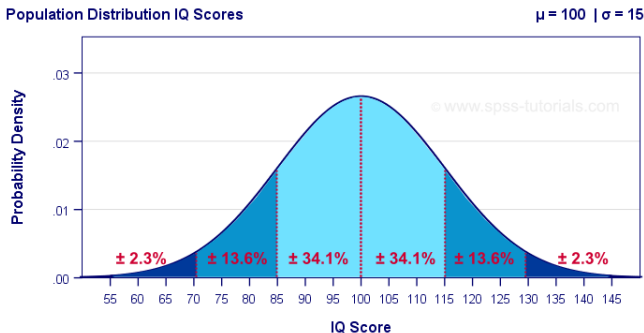
SOME CONTINUOUS PROBABILITY DISTRIBUTION

I. Normal distribution



Normal distribution (Gauss's distribution) is one of the most important probability distributions in statistical theory. The natural world as well as many economic and social laws obey this law of normal distribution.

Example. random quantities follow the law of normal distribution such as: IQ, height, chest circumference, weight of human, length of human sleep, fluctuations in stock value,



Hình: IQ score

DEFINITION

The random variable X is said to have a normal distribution, denoted by $X \sim N(\mu, \sigma^2)$, if its density function is determined by

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

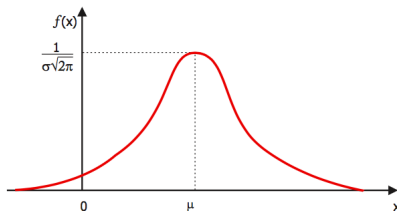
where μ — the mean of X , σ — the standard deviation of X .

DEFINITION

The random variable X is said to have a normal distribution, denoted by $X \sim N(\mu, \sigma^2)$, if its density function is determined by

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

where μ — the mean of X , σ — the standard deviation of X .



The curve of the density function is bell-shaped, symmetrical about the line $x = \mu$ and takes Ox as the horizontal asymptote. The peak of the graph of the function $f(x)$ reaches at

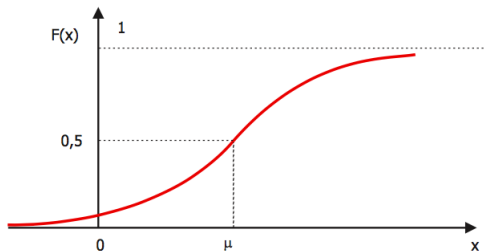
$$\max f(x) = f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}.$$

The cumulative function of X :

$$F(x) = \int_{-\infty}^x f(t)dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

The cumulative function of X :

$$F(x) = \int_{-\infty}^x f(t) dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$



The curve of the graph of $F(x)$ has a left horizontal asymptote with the Ox axis, a right horizontal asymptote with $y = 1$, symmetric about the point $(\mu; 0,5)$.

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

$$\begin{aligned} EX &= \mu, & DX &= \sigma^2, \\ \text{mod } X &= \text{median } X = \mu. \end{aligned}$$

Standard normal distribution

The distribution of a normal random variable with mean 0 and variance 1 is called a standard normal distribution.

A random variable Z is said to have a standard normal distribution If it has a normal distribution $Z \sim N(0, 1)$, in this case the density function of Z :

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

The cumulative function of Z , denoted by $\Phi(x)$, is determined by

$$\Phi(x) = \int_{-\infty}^x f(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

Notice.

- You can look up the values of $\Phi(x)$ in the pre-calculated table.
- The pre-calculated table usually only gives values of $\Phi(x)$ for $x > 0$. If $x < 0$, you can use the following formula:

$$\Phi(-x) = 1 - \Phi(x).$$

- $P\{a < Z < b\} = \Phi(b) - \Phi(a)$, $P\{Z > a\} = 1 - P\{Z < a\} = 1 - \Phi(a)$.
- $EZ = 0$, $DZ = 1$, $\text{mod } Z = \text{median } Z = 0$.

TRANSFORMATION OF VARIABLES

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

For convenience in calculation, we reduce the normal distribution to a normal distribution by transforming variables:

$$Z = \frac{X - \mu}{\sigma}, \quad \text{hay} \quad X = \sigma Z + \mu.$$

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

PROBABILITY FORMULAS

Consider a random variable $X \sim N(\mu, \sigma^2)$.

We can calculate the probabilities related to X by bringing in the quantity Z with a standard normal distribution and then looking up the table.

$$P\{X < a\} = P\{\sigma Z + \mu < a\} = P\{Z < \frac{a - \mu}{\sigma}\} = \Phi(\frac{a - \mu}{\sigma}),$$

$$P\{X > a\} = 1 - P\{X < a\} = 1 - \Phi(\frac{a - \mu}{\sigma}),$$

$$P\{a < X < b\} = P\{\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\} = \Phi(\frac{b - \mu}{\sigma}) - \Phi(\frac{a - \mu}{\sigma}).$$

Example 1. Suppose X has a normal distribution with parameters $\mu = 2000$ và $\sigma = 100$.

Given $\Phi(0,5) = 0,6915$, $\Phi(2) = 0,9772$, $\Phi(1,5) = 0,9332$, $\Phi(1,881) = 0,97$.

Find:

- $P\{X > 2150\}$,
- $P\{1800 < X < 2050\}$,
- a if $P\{X > a\} = 0,03$.

Example 1. Suppose X has a normal distribution with parameters $\mu = 2000$ và $\sigma = 100$.

Given $\Phi(0,5) = 0,6915$, $\Phi(2) = 0,9772$, $\Phi(1,5) = 0,9332$, $\Phi(1,881) = 0,97$.

Find:

- $P\{X > 2150\}$,
- $P\{1800 < X < 2050\}$,
- a if $P\{X > a\} = 0,03$.

Solution.

- $P\{X > 2150\} = 1 - P\{X < 2150\} = 1 - \Phi\left(\frac{2150-2000}{100}\right) = 1 - \Phi(1,5) = 1 - 0,9332 = 0,0668$.
- Following the formula
$$P\{a < X < b\} = P\left\{\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$
therefore
$$P\{1800 < X < 2050\} = \Phi\left(\frac{2050-2000}{100}\right) - \Phi\left(\frac{1800-2000}{100}\right) = \Phi(0,5) - \Phi(-2) = 0,6915 - (1 - 0,9772) = 0,6687$$
.

Example 1. Suppose X has a normal distribution with parameters $\mu = 2000$ và $\sigma = 100$.

Given $\Phi(0,5) = 0,6915$, $\Phi(2) = 0,9772$, $\Phi(1,5) = 0,9332$, $\Phi(1,881) = 0,97$.

Find:

- $P\{X > 2150\}$,
- $P\{1800 < X < 2050\}$,
- a if $P\{X > a\} = 0,03$.

Solution.

- $P\{X > 2150\} = 1 - P\{X < 2150\} = 1 - \Phi\left(\frac{2150-2000}{100}\right) = 1 - \Phi(1,5) = 1 - 0,9332 = 0,0668$.
- Following the formula
$$P\{a < X < b\} = P\left\{\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$
therefore
$$P\{1800 < X < 2050\} = \Phi\left(\frac{2050-2000}{100}\right) - \Phi\left(\frac{1800-2000}{100}\right) = \Phi(0,5) - \Phi(-2) = 0,6915 - (1 - 0,9772) = 0,6687$$
.
- $P\{X > a\} = 1 - \Phi\left(\frac{a-2000}{100}\right) = 0,03$.
 $\implies \Phi\left(\frac{a-2000}{100}\right) = 0,97$.
Since $\Phi(1,881) = 0,97$, therefore $a = 2000 + 100 \cdot 1,881 = 2188,1$.

Example 2. Assume the weight of a cookie package is normally distributed. In 1000 packages of this type of cookie, there are 70 packages weighing more than 1015g. Calculate how many packages of the cookie weigh less than 1008g, knowing that the average weight of 1000 packages of the cookie is 1012g.

Example 2. Assume the weight of a cookie package is normally distributed. In 1000 packages of this type of cookie, there are 70 packages weighing more than 1015g. Calculate how many packages of the cookie weigh less than 1008g, knowing that the average weight of 1000 packages of the cookie is 1012g.

Solution. By assumption we have $\mu = 1012$. Let X be the weight of a cookie package. Then X has a normal distribution with unknown $\mu = 1012$ and σ^2 . We will find σ .

Since $P\{X > 1015\} = 70 : 1000 = 0,07$, therefore

$$\begin{aligned}P\{X > 1015\} &= 1 - \Phi\left(\frac{1015 - 1012}{\sigma}\right) = 0,07 \\ \implies \Phi(3/\sigma) &= 0,93 = \Phi(1,476) \implies \sigma = 2,0325.\end{aligned}$$

From there we reach

$$P\{X < 1008\} = \Phi\left(\frac{1008 - 1012}{2,0325}\right) = \Phi(-1,968) = 0,0245.$$

So, it can be deduced that in 1000 packages of the cookie there will be: $1000 \cdot 0,0245 = 24,5$ packages weighing less than 1008g.

THANK YOU FOR YOUR ATTENTION!