STAT 320: Principles of Probability

Unit 6 Part:C

The Normal Distribution

United Arab Emirates University

Department of Statistics

Normal Distribution

The normal distribution is one of the most commonly used probability distribution for applications:

- When we repeat an experiment numerous times and average our results, the random variable representing the average or total tends to have a normal distribution as the number of experiments becomes large.
- The previous fact, which is known as the central limit theorem, is fundamental to many of the statistical techniques we will discuss later.
- Many physical characteristics (Heights, weights, etc.) tend to follow a normal distribution.
- Errors in measurement or production processes can often be approximated by a normal distribution.
- Under certain conditions, many probability distributions can be approximated by a normal distribution.

The Normal Distribution denoted by Normal(μ, σ^2) is characterized by two parameters, namely the mean $\mu \in \mathbb{R}$ and the standard deviation $\sigma > 0$.

Definition (Normal Distribution)

A continuous random variable X is said to be normally distributed with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, if its probability density function is given as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for} \quad -\infty < x < \infty.$$

If $X \sim$ Normal(μ, σ^2) then $E(X) = \mu$ and , $\mathrm{Var}(X) = \sigma^2$

Let $X \sim \text{Normal}(\mu, \sigma^2)$ for $\mu \in \mathbb{R}, \sigma^2 > 0$

Mean

$E(X) = \mu$

Variance

$\text{VAR}(X) = \sigma^2$

MGF

$M_X(t) = e^{\mu t + \frac{t^2\sigma^2}{2}}$

Distribution	Support $\mathbb{S}_X$	pdf $f_X(x)$	Mean $E(X)$	Variance $\text{Var}(X)$	mgf $M_X(t)$
$\text{Normal}(\mu, \sigma^2)$ mean $= \mu$, Variance $= \sigma^2$	$(-\infty, \infty)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for $x \in \mathbb{R}$	μ	σ^2	$e^{\mu t + \frac{t^2\sigma^2}{2}}$

Standard Normal Normal$(0,1)$,(i.e. Normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$,) is referred to as the **Standard Normal** Distribution.

Z-Transformation If $X \sim$ Normal(μ, σ^2) for some $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ then ,

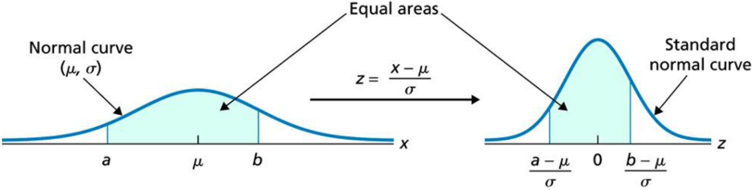
$$\boxed{Z \sim \text{Normal}(0,1) \text{ where } Z = \frac{X-\mu}{\sigma}}$$

The role of the parameters μ and σ^2

Finding Probabilities Using Normal CDF

The cdf of the normal distribution, does not have a closed form analytical expression and the corresponding integral is nontrivial to evaluate.

The same is true for the cdf of the standard normal distribution (commonly denoted by $\Phi(z)$). However, $\Phi(z)$ is usually tabulated for values of z from -3.49 to 3.49 in increments of 0.01 and can be usedto calculate any normal probability by standardizing it first.



Example

Example : The times of first failure of a unit of a brand of ink jet printers are approximately normally distributed with a mean of 1,500 hours and a standard deviation of 200 hours. Use the statistical calculator.

a). What fraction of these printers will fail before 1,000 hours?

b). What is the probability that the first failure time of a selected printer will fail be between 1,300 and 1700 hours?

Example

$X \sim \text{Normal}(\mu = 1500, \sigma^2 = 200^2)$.

$$
\begin{aligned}
& P(X < 1000) \\
= \; & P\left(\frac{X-\mu}{\sigma} < \frac{1000-\mu}{\sigma}\right) \\
= \; & P\left(Z < \frac{1000-1500}{200}\right) \\
= \; & P\left(Z < -2.5\right) \\
= \; & \Phi(-2.5) \\
= \; & 0.0062
\end{aligned}
$$

Example

$X \sim \text{Normal}(\mu = 1500, \sigma^2 = 200^2)$.

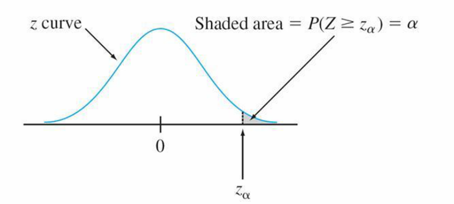
$$
\begin{aligned}
& P(1300 < X < 1700) \\
= \; & P\left(\frac{1300 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{1700 - \mu}{\sigma}\right) \\
= \; & P\left(\frac{1300 - 1500}{200} < Z < \frac{1700 - 1500}{200}\right) \\
= \; & P\left(-1 < Z < -1\right) \\
= \; & \Phi(1) - \Phi(-1) \\
= \; & 0.8413 - 0.1587 \\
= \; & 0.6826
\end{aligned}
$$

Backward Normal calculations and Percentiles

We could find the observed value (x) of a given proportion or percentile in Normal(μ, σ^2) by unstandardizing the z-value as follows:

1. Find the z-value corresponding to the lower tail probability using $\Phi^{-1}(\cdot)$
2. Unstandardize $x = \mu + \sigma z$.

In the standard normal distribution, z will denote the z-value for which of the area under the standard normal curve lies to the right of z, i.e. $P(Z \geq z_\alpha)$



Example

Example : The times of first failure of a unit of a brand of ink jet printers are approximately normally distributed with a mean of 1,500 hours and a standard deviation of 200 hours. Use the statistical calculator.

a). what should be the guarantee time for these printers if the manufacturer wants only 5% to fail within the guarantee period.

$X \sim \text{Normal}(\mu = 1500, \sigma^2 = 200^2)$. Now we want to find a such that $P(X < a) = 0.05$, so

$$
\begin{aligned}
& P(X < a) = 0.05 \\
\Longrightarrow\quad & P\left(\frac{X-\mu}{\sigma} < \frac{a-\mu}{\sigma}\right) = 0.05 \\
\Longrightarrow\quad & P\left(Z < \frac{a-\mu}{\sigma}\right) = 0.05 \\
\Longrightarrow\quad & \Phi\left(\frac{a-\mu}{\sigma}\right) = 0.05 \\
\Longrightarrow\quad & \frac{a-\mu}{\sigma} = \Phi^{-1}(0.05) \\
\Longrightarrow\quad & a = \mu + \sigma\Phi^{-1}(0.05) \\
\Longrightarrow\quad & a = 1500 + 200 \times (-1.64) \\
=\quad & 1172.
\end{aligned}
$$

Exercise

Example : An engineer working for a manufacturer of electronic components takes a large number of measurements of a particular dimension of components from the production line. She finds that the distribution of dimensions is normal, with a mean of 2.340 cm and a standard deviation of 0.06 cm.

a). What percentage of measurements will be less than 2.45 cm?

b). What percentage of dimensions will be between 2.25 cm and 2.45 cm?

c). What value of the dimension will be exceeded by 98% of the components?

Exercise

Example : Wires manufactured for use in a computer system are specified to have resistances between 0.12 and 0.14 ohms. The actual measured resistances of the wires produced by company A have a normal probability distribution with mean 0.13 ohm and standard deviation 0.005 ohm.

a). What is the probability that a randomly selected wire from company A's production will meet the specifications?

b). If four of these wires are used in each computer system and all are selected from company A, what is the probability that all four in a randomly selected system will meet the specifications?

Exercise

Example : The SAT and ACT college entrance exams are taken by thousands of students each year. The mathematics portions of each of these exams produce scores that are approximately normally distributed. In recent years, SAT mathematics exam scores have averaged 480 with standard deviation 100. The average and standard deviation for ACT mathematics scores are 18 and 6, respectively.

a). An engineering school sets 550 as the minimum SAT math score for new students. What percentage of students will score below 550 in a typical year?

b). What score should the engineering school set as a comparable standard on the ACT math test?

Exercise

Example : Of the Type A electrical resistors produced by a factory, 85% have resistance greater than 41 ohms, and 3.7% of them have resistance greater than 45 ohms. The resistances follow a normal distribution.

a). What percentage of these resistors have resistance greater than 44 ohms?

Outline

Moment Generating Function

Questions?