10.10: Common Probability Distribution

Prof. Josefina V. Almeda

Stat 115

Normal Distribution

Definition 10.27

A continuous random variable X is said to be **normally distributed** if its probability density function is given by :

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

for any real number x. The constants, μ and σ^2, are such that $-\infty < \mu < \infty$ and $\sigma^2 > 0$. The values, e and π, are mathematical constants, wherein, $e \approx 2.71828$ and $\pi \approx 3.14159$.

Characteristics of Normal Distribution

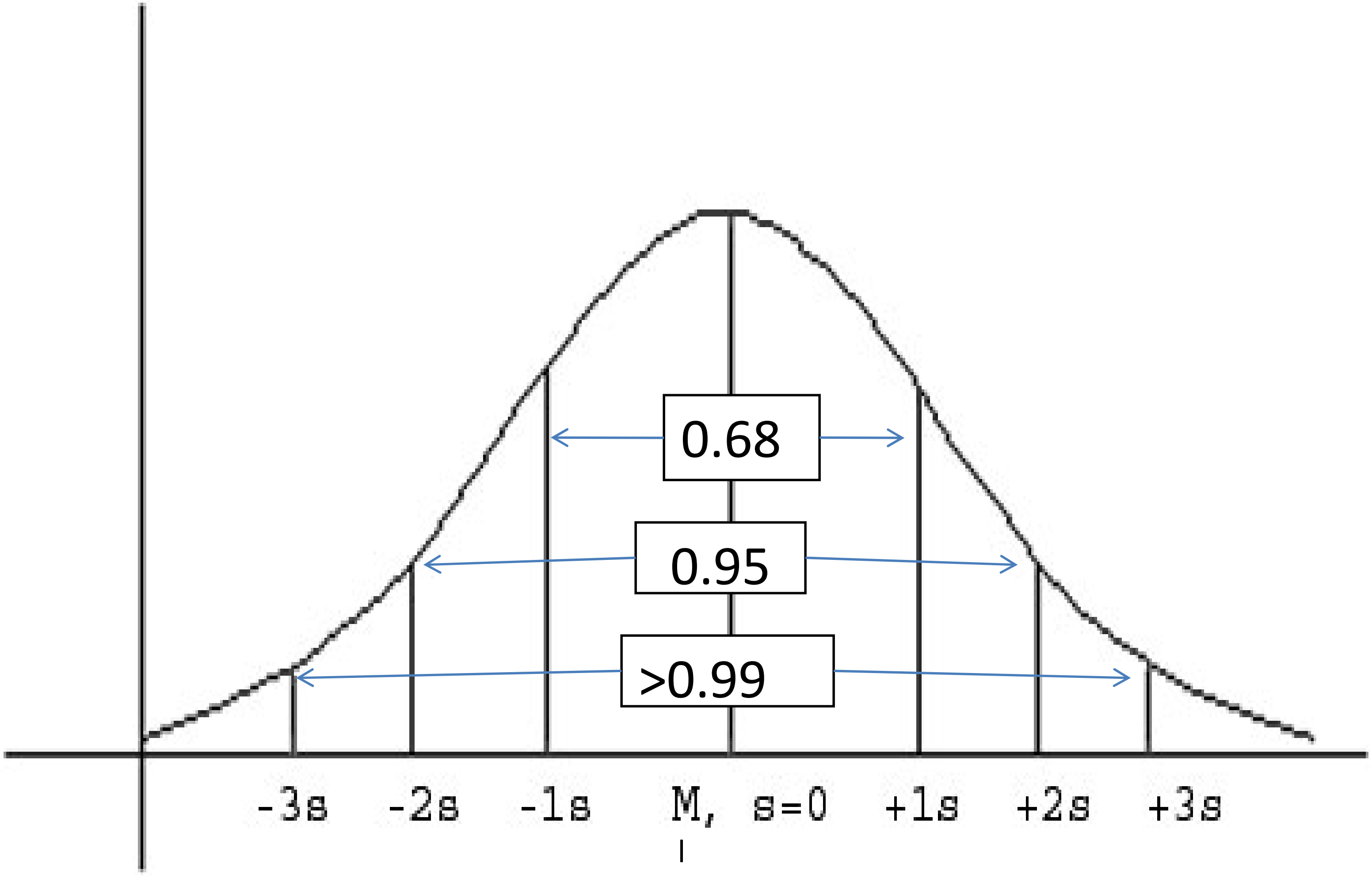
- Bell-shaped curve
- Total area bounded by this curve and the x-axis is equal to 1
- Curve is symmetric about the constant μ
- Curve will approach the x-axis as we proceed in either direction away from μ, but it will never touch the x-axis.

Normal Distribution Properties

For a normal curve, the area within:

a) one standard deviation from the mean is about 68%,
b) two standard deviations from the mean is about 95%; and
c) three standard deviations from the mean is about 99.7%.

FIGURE 10.10.



NORMAL CURVE SHOWING THE PROBABILITIES IN THE DIFFERENT REGIONS

Definition 10.28

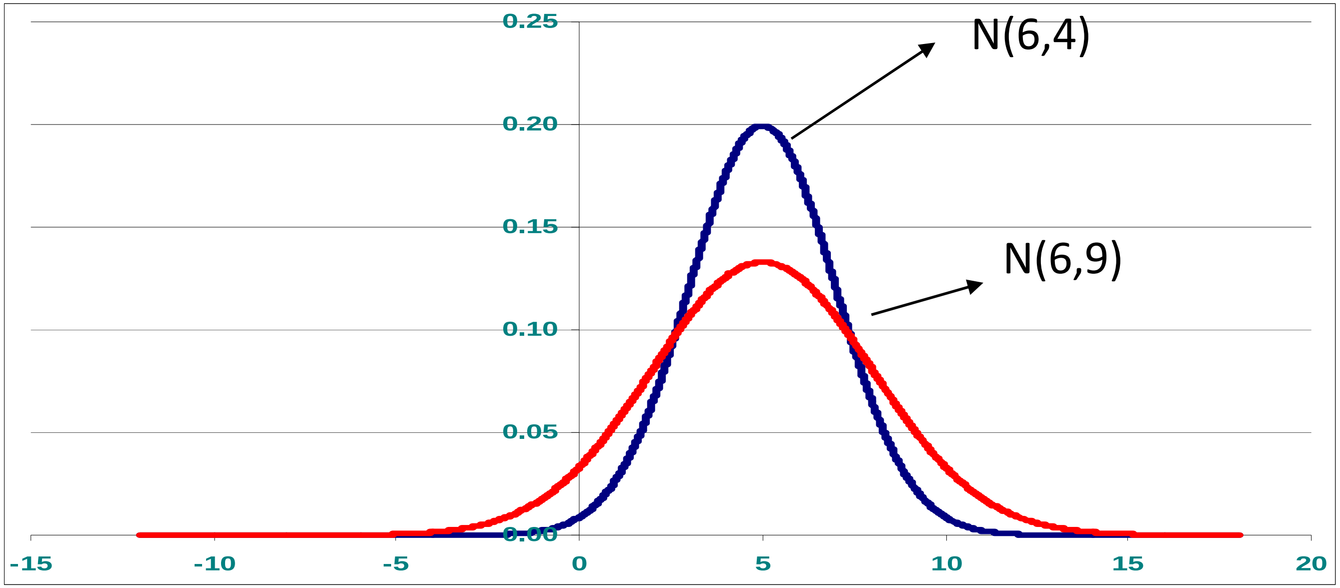
A parameter is a constant that determines the specific form of the probability distribution.

Theorem 10.14

If $X \sim Normal(\mu, \sigma^2)$ then $E(X) = \mu$ and $Var(X) = \sigma^2$

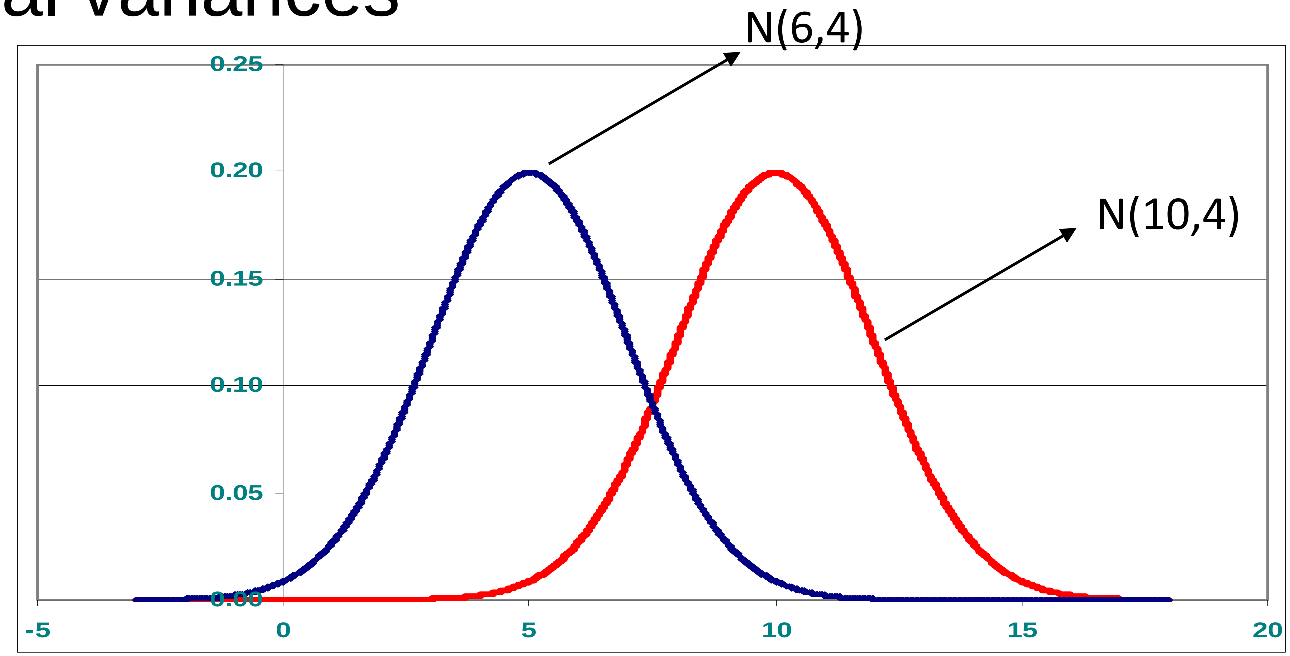
THE NORMAL CURVE

Two normal distributions with the same mean but different variances.



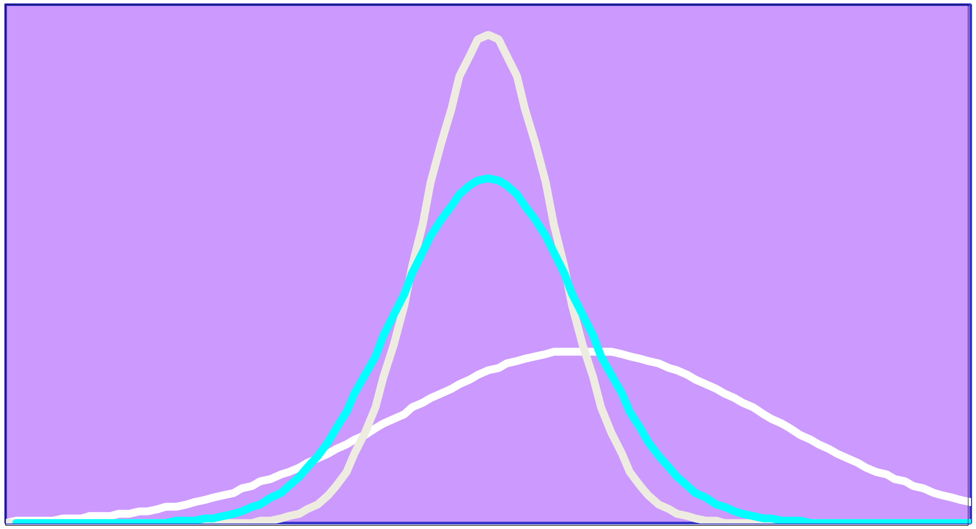
THE NORMAL CURVE

Two normal distributions with the different means but equal variances



Many Normal Distributions

There are an infinite number of normal curves



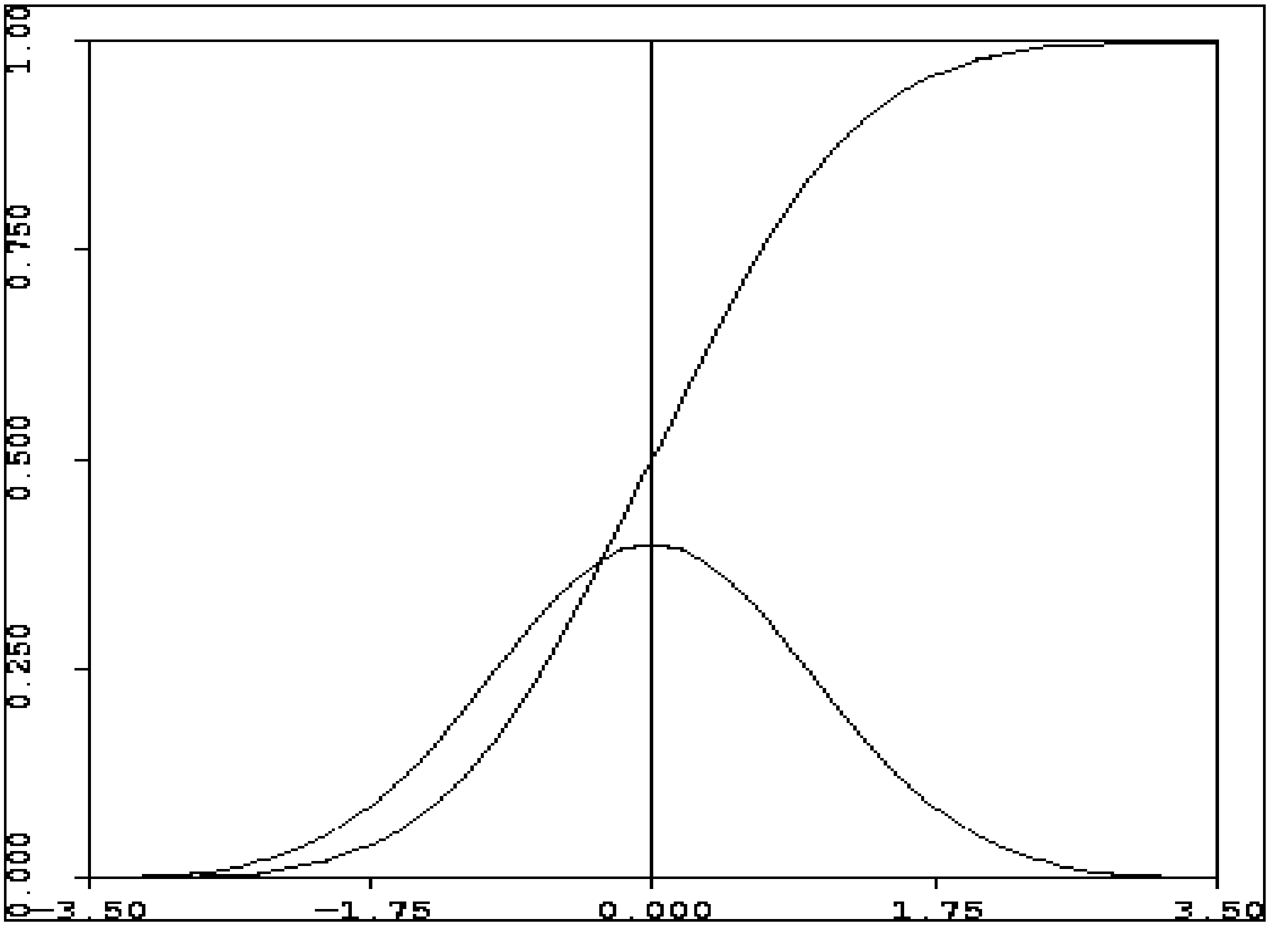
By varying the parameters σ and μ, we obtain different normal distributions

Definition 10.29: Standard normal random variable

If the normal random variable has mean 0 and variance 1, it is called a ***standard normal random variable*** and is denoted by Z.

NOTES: Standard normal random variable

- If Z is a standard normal random variable then we write Z~Normal(0,1).
- Figure 10.12 shows the graph of its pdf and cdf. The bell-shaped curve is the pdf while the S-shaped curve is its cdf.
- As expected, its cdf is a continuous function since Z is a continuous random variable.
- Table B.1 in Appendix B, presents the values of the cdf of Z at certain points ranging from -3.49 to 3.49. We can use this table to evaluate the probability of any event expressed in terms of Z.


1.00
0.750
0.500
0.250
0.000
-3.50
-1.75
0.000
1.75
3.50

Example 10.49

Suppose Z~N(0,1). Use Table B.1 in Appendix B, to determine the $P(Z \leq 2.54)$ and $P(Z. \leq -1.27)$

Solution: To determine the $P(Z \leq 2.54)$, we proceed down the first column (the z-column) to the entry "2.5". Then we move across the table until we reach the entry that falls in the column labeled "0.04" at the top of the table to read the value of $P(Z \leq 2.54)$ as 0.9945.

Formulas to be Used

1. $P(Z \leq a) = P(Z < a) = F(a)$.
2. $P(Z > a) = P(Z \geq a) = 1 - F(a)$.
3. $P(a < Z < b) = P(a \leq Z \leq b) = P(a \leq Z < b)$

   $= P(a < Z \leq b) = F(b) - F(a)$.

F(a) in these formulas refers to the CDF of the standard normal random variable evaluated at the point a.

Example 10.50:

Suppose $Z \sim N(0,1)$. Determine the following probabilities:

a) $P(-1.75<Z<1.62) = F(1.62) - F(-1.75)$

$= 0.9474 - 0.0401 = 0.9073$

b) $P(Z > 0.75) = 1 - F(0.75) = 1 - 0.7734 = 0.2266$.

Example

The number of calories in a salad on the lunch menu can be approximated by the normal distribution with mean=200 and sd=5. Find the probability that the salad you select will contain

a) More than 208 calories

b) Between 190 and 200 calories

Exercises

1. Given a normal distribution with $\mu = 40$ and
$\sigma = 8$, find the probability that X assumes a value
a. Less than 45
b. Between 35 and 45
c. More than 45
2. Given the normally distributed random variable X with mean 18 and standard deviation 2.4, find
a. The value k such that $P(X < k) = 0.2578$
b. The value k such that $P(X > k) = 0.1539$

Exercises

3. A softdrink machine is regulated so that it dispenses an average of 200 ml per cup. IF the amount of drink dispensed is normally distributed with a standard deviation equal to 15 ml,

a. What fraction of the cups will contain more than 224 ml?

b. What is the probability that a cup contains between 191 ml and 209 ml?

c. How many cups will likely overflow if 230 ml cups are used for the next 1000 drinks?

d. Below what value do we get the smallest 25% of the drinks?

Binomial Distribution

Definition 10.31

A **binomial experiment** is a random experiment that satisfies the following properties:

a) It consists of observing the outcomes of a sequence of *n* trials.

b) Each trial can result in one of only 2 possible outcomes which we can label as S-for success and F-for failure.

c) The probability of success must be the same for each one of the *n* trials. We will denote this by *p*.

d) The trials are independent in the sense that the probability of success at a particular trial should not be affected by the outcomes of the previous trials.

Definition 10.32: Binomial Distribution

A discrete random variable X is said to follow a **binomial distribution** if its probability mass function is given by:

$$P(X=x)=\binom{n}{x}p^{x}(1-p)^{n-x}, \quad x=0,1,2,\ldots n$$

where n and p are such that n is a positive integer and p is any real number between 0 and 1.

Notes: Binomial Distribution

- If X follows the binomial distribution, then we write X~ Bi(n,p)
- The binomial distribution has 2 parameters:

  n = number of trials

  p = probability of the event that the outcome of the trial is a success.

Theorem 10.16

If $X \sim Bi(n,p)$ then $E(X)=np$ and $Var(X)=npq$, where $q=1-p$.

Illustration: Binomial Distribution*

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that exactly 5 survive.

Solution: Let X be the number of people that survive.

P(X=5) = b(5;15, 0.4) = $\binom{15}{5} 0.4^5 (1-0.4)^{10}$ = 0.1859

*from Walpole, Introduction to Statistics

Exercises

1. A multiple choice quiz has 15 questions, each with 4 possible answers of which only 1 is the correct answer. What is the probability that sheer guesswork yields
   a. Exactly 10 correct answers
   b. At least 1 correct answer
   c. 10 to 12 correct answers.
2. Suppose that airplane engines operate independently in flight and fail with probability 1/5. Assuming that a plane make a safe flight if at least one-half of its engines run, which between a 4-engine plane and a 2-engine plane has the higher probability for a successful flight?