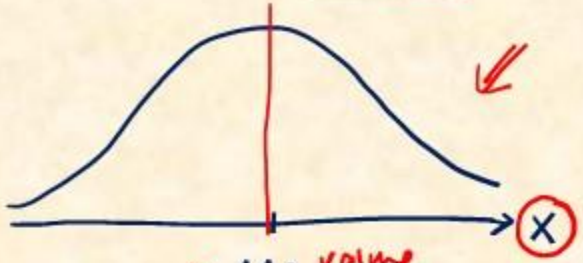
Normal Dist.

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



$\mu =$ value

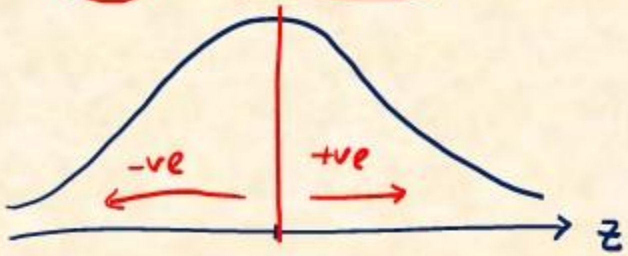
$\sigma =$ value

$$Z = \frac{X-\mu}{\sigma}$$

$\Longrightarrow$

Standard Normal Dist

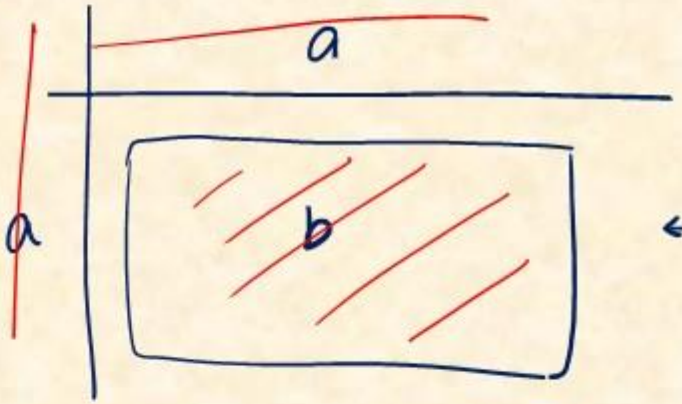
$Z \sim N(0, 1)$

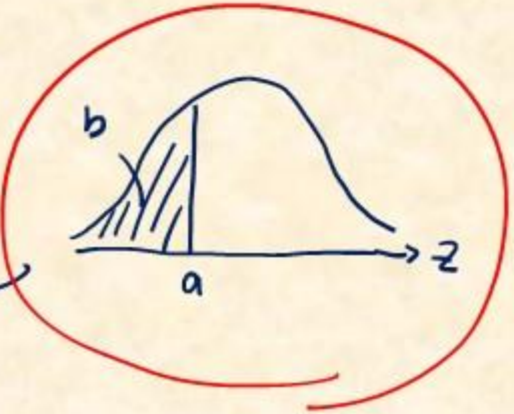


$\mu = 0$

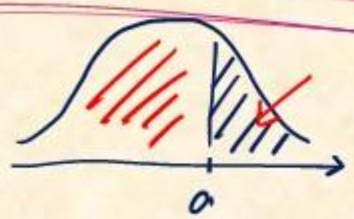
$\sigma = 1$

$P(z < a) = b$



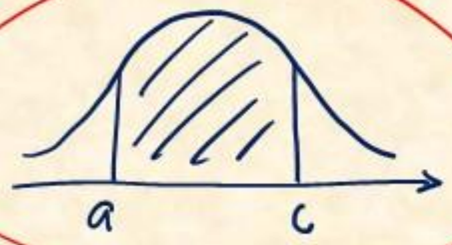


$P(z > a)$



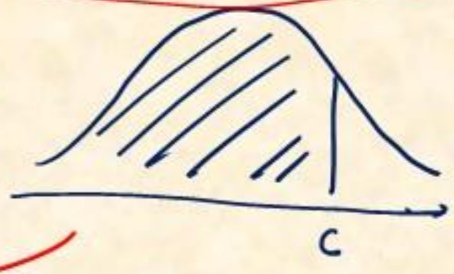
$= 1 - P(z < a)$

$P(a < z < c)$

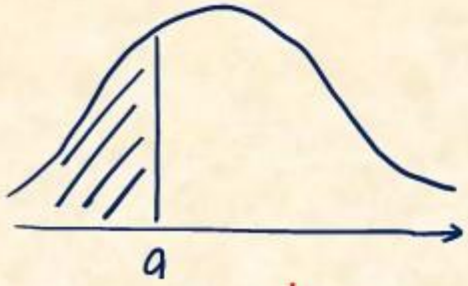


$= P(z < c) - P(z < a)$

=



−



Example 7.4-4

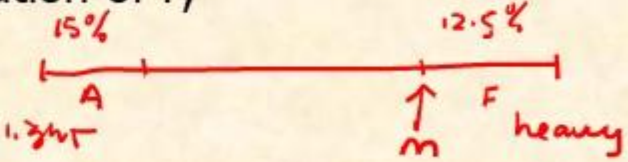
The weights of a certain batch of obese male recruits are approximately normally distributed with mean 88 kg and standard deviation 9. The lightest 15% of the recruits receive a classification of A whilst the heaviest 12.5% receive a classification of F. Find

(i) the minimum weight required to obtain a classification of F,

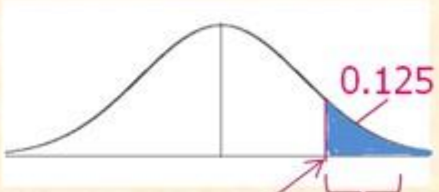
Let X be the weight of an obese male recruit

$$\Rightarrow X \sim N(88, 9^2)$$

$\mu \quad \sigma^2$



(i) Let m be the minimum weight to be in classification F.



$$P(X \geq m) = 0.125 \qquad 1 - P(X < m) = 0.125$$

$$P(X < m) = 1 - 0.125 = 0.875$$

$$Z = \frac{X - \mu}{\sigma}$$

$$P\left(Z < \frac{m - 88}{9}\right) = 0.875$$

a b

From standard normal table, $z = 1.15$

$$\frac{m - 88}{9} = 1.15$$

$$m = 1.15(9) + 88 = 98.35 \text{ kg}$$

z		0.05	0.06
1.1		0.8749	0.8770

closest to 0.875

Example 7.4-4

μ σ

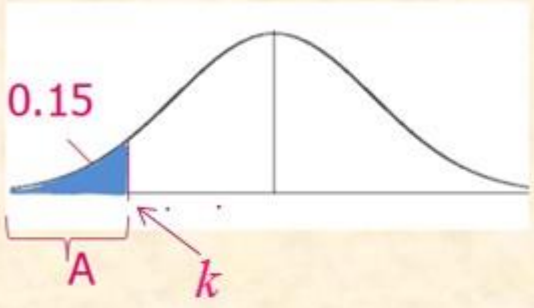
The weights of a certain batch of obese male recruits are approximately normally distributed with mean 88 kg and standard deviation 9. The lightest 15% of the recruits receive a classification of A whilst the heaviest 12.5% receive a classification of F. Find

(ii) the weight of the heaviest recruit in classification A.

A k

Let X be the weight of an obese male recruit

$$X \sim N(88, 9^2)$$



(ii) Let k be the largest weight to be in classification A.

$$P(X \le k) = 0.15$$

$$P\left(Z < \frac{k-88}{9}\right) = 0.15$$

From standard normal table, $z = -1.04$

$$\frac{k-88}{9} = -1.04$$

$$k = -1.04(9) + 88 = 78.64 \text{ kg}$$

z		0.04	0.05
-1.0		0.1492	0.1515

closest to 0.15

Computing Mathematics 2

Chapter 8 : Distribution of Sample Means

Objectives:

At the end of this lesson, the student should be able to:

sampling dist.

- identify distribution of sample means
- apply the Central Limit Theorem to find the probability of a sample mean for sufficiently large samples

8.1 Introduction

In Chapter 1 → μ
we introduced the concept of sample mean, $\bar{x}$, for **a single sample data set**.
In this case, $\bar{x}$ is a single value.

In Chapter 8 →
we will look at the sample means of **multiple samples** and the **distribution of the sample means**.

Introduction example

- Let X denote the weight of a single peanut from a packet of peanuts.
- Suppose we weigh each peanut in that packet.
- We can calculate the <u>sample mean</u> of the weight of the peanuts, $\bar{x}$, in that single packet.

1 packet:

- Suppose we have *n* packets of peanuts. We will calculate the sample mean of each packet of peanuts.
- The <u>sample mean for a packet of peanut</u> will most likely be *different from* the <u>sample mean of other packets</u>.

- Hence, in general, the sample mean, $\overline{X}$, is a **<u>random variable</u>** (as we cannot determine the actual value of $\bar{x}$ for a randomly chosen sample.)

8.1 Introduction $\overline{X}$

- Since $\overline{X}$ is a random variable, we can calculate probabilities involving X once we know its distribution which is shown below:
- Let μ and σ^2 be the mean and variance of a random variable X (single quantity). If we have a sample of n objects,
- by Central Limit Theorem,

chap 8

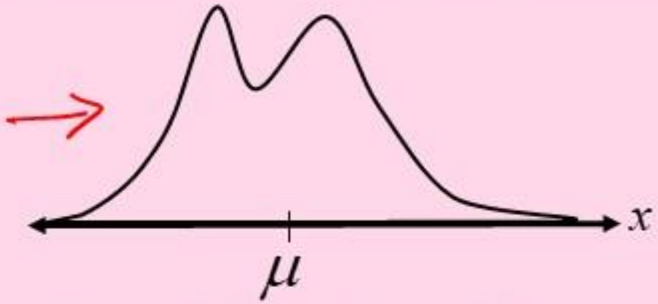
(a) If $X \sim N(\mu, \sigma^2)$, then $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

(b) If the distribution of X is **not normal** (or unknown), and $n \geq 30$

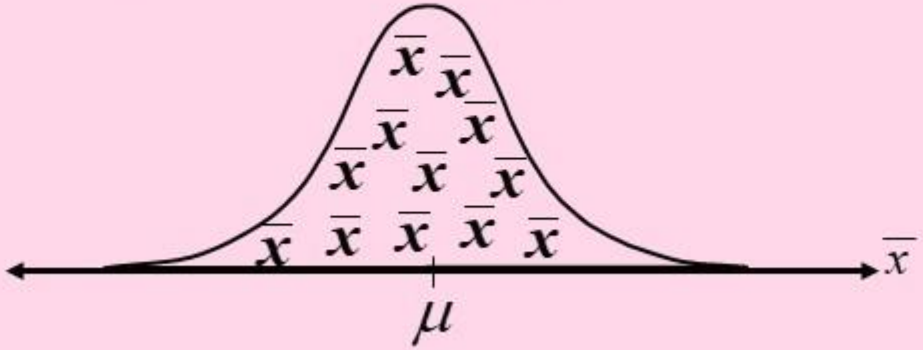
$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$ **approximately**.

The central limit theorem

1. If samples of size $n \geq 30$, are drawn from any population with mean = μ & standard deviation = σ,

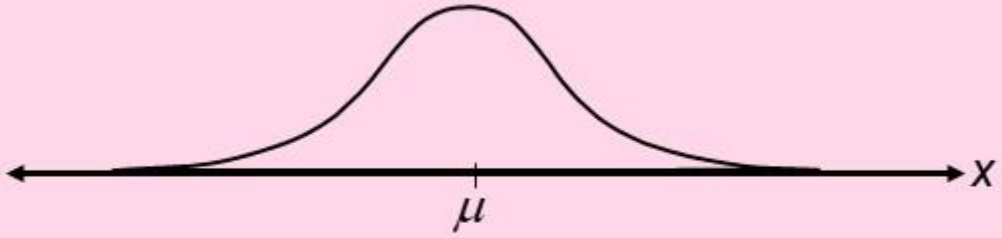


then the sampling distribution of the sample means approximates a normal distribution. The greater the sample size, the better the approximation.

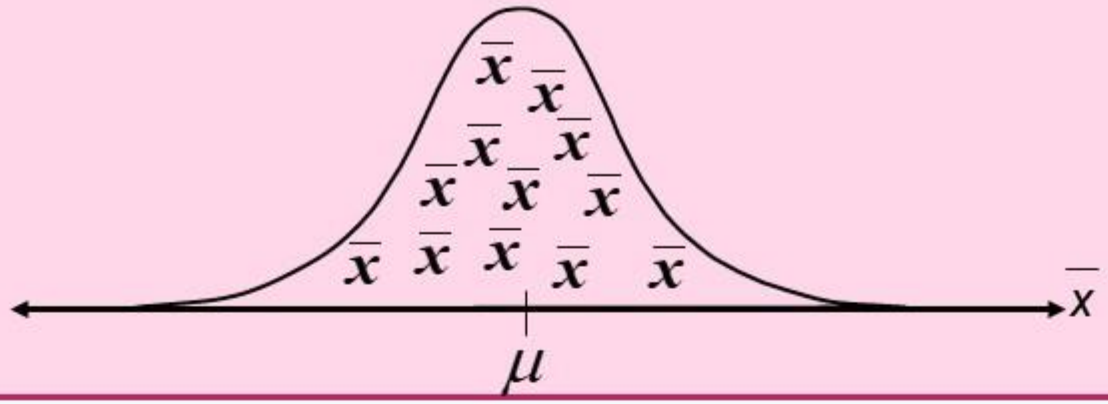


The central limit theorem

2. If the population itself is normally distributed,



the sampling distribution of the sample means is normally distribution for ***any*** sample size n.



Central limit theorem

population is normal dist.

sample $n \geq 30$
or
$n < 30$
→ normal dist.

Since population is normal dist,
∴ sample is also normal dist
(by central limit theorem)

population is NOT normal dist

sample $n \geq 30$
→ normal Dist

sample $n < 30$
→ NOT normal Dist

Even though population is not normal dist, but the sample size is ≥ 30, ∴ sample is approx normal dist (by central limit theorem)

Normal Dist

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

$\downarrow \quad z = \dfrac{X-\mu}{\sigma}$

$z \sim N(0,1)$

Dist of sample mean
(sampling Dist)

$\bar{X} \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$

$\dfrac{\sigma^2}{n}$

$\downarrow \quad z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

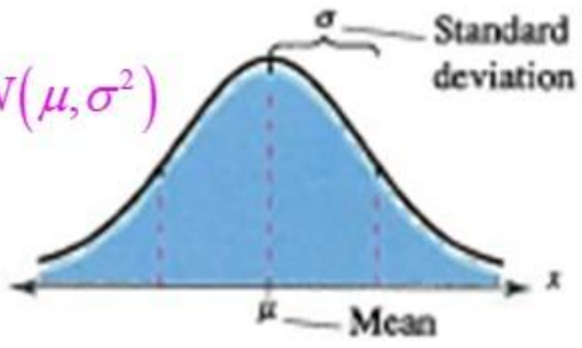
$\sqrt{\dfrac{\sigma^2}{n}} = \dfrac{\sigma}{\sqrt{n}}$

$z \sim N(0, 1)$

Summary of Central Limit Theorem

1. Population Distribution - Normal

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



Distribution of Sample Means, (any n)

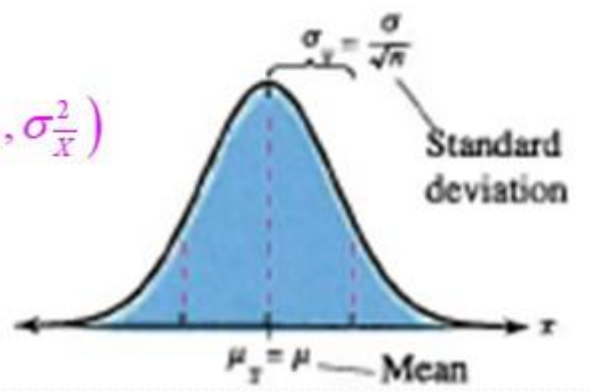
$\overline{X} \sim N\left(\mu_{\overline{X}}, \sigma^2_{\overline{X}}\right)$

where

$\mu_{\overline{X}} = \mu$

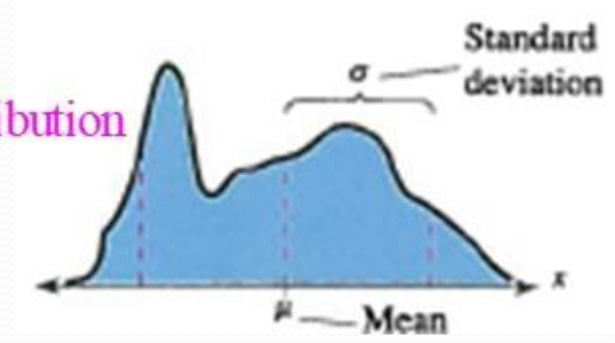
$\sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}$

$\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$ = standard error of the mean



2. Population Distribution - Not Normal

$X \sim$ any distribution



Distribution of Sample Means, ($n \geq 30$)

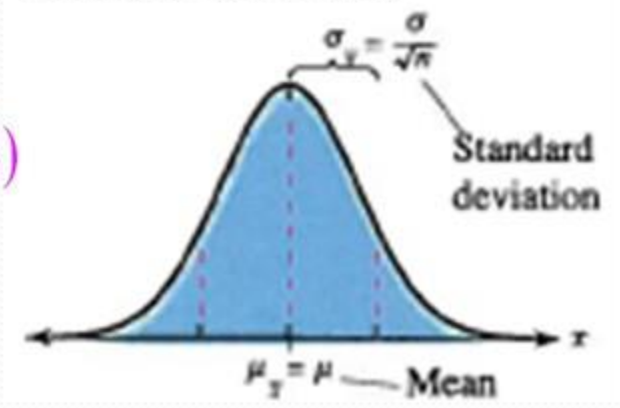
$\overline{X} \sim N\left(\mu_{\overline{X}}, \sigma^2_{\overline{X}}\right)$

where

$\mu_{\overline{X}} = \mu$

$\sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}$

$\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$ = standard error of the mean



Central Limit Theorem

In either case,

Mean of the sample mean, $\overline{X}$ $\mu_{\overline{X}} = \mu$

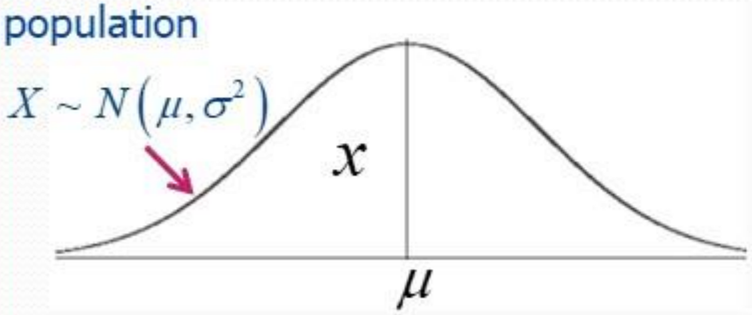
Variance of the sample mean, $\overline{X}$ $\sigma_{\overline{X}}^2 = \dfrac{\sigma^2}{n}$

std dev of the dist of sample mean.

Standard error of the mean $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$

Difference between transforming

X to z-score



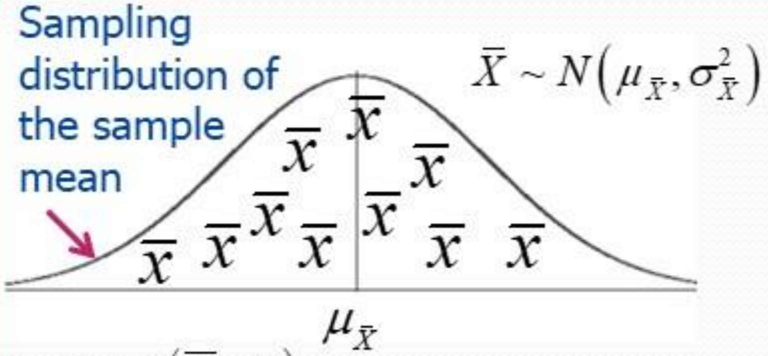
To find P(X < 2):

$$\therefore P\left(X < 2\right)$$

$$= P\left(Z < \frac{2-\mu}{\sigma}\right)$$

$$Z = \frac{X - \mu}{\sigma}$$

$\overline{X}$ to z-score



To find $P\left(\overline{X} < 2\right)$:

$$\therefore P\left(\overline{X} < 2\right)$$

$$= P\left(Z < \frac{2-\mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$Z = \frac{\overline{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

where

$$\mu_{\bar{x}} = \mu;$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Example 8.2-1

σ pop is nor dist μ

The mass of garlic bulbs produced by a particular farm is approximately normally distributed, with a mean of 60g and a standard deviation of 5g. State the distribution of the sample mean of a random sample of 16 garlic bulbs.

n

Solution:

Let X be the mass of a garlic bulb. $X \sim N\left(60, 5^2\right)$

$\frac{\sigma^2}{n}$

The sample mean of 16 garlic bulbs, $\bar{X} \sim N\left(\underset{\mu}{60}, \frac{25}{16}\right)$ by CLT.

$$where$$

$$\mu_{\bar{X}} = \mu = 60$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} = \frac{5^2}{16} = \frac{25}{16}$$

Example 8.2-2

Not Normal Dist

The waistline of forty-year-old male Singaporeans is known to have a mean of 33 (μ) inches and a variance of 9 (σ^2) inches. A random sample of 36 (n) forty-year-old male Singaporeans was selected. Find the probability that the sample mean

(a) is greater than 31.5 inches, $P(\bar{x} > 31.5)$

(b) lies between 32 and 34 inches, $P(32 < \bar{x} < 34)$

(c) differs from the population mean by more than one inch.

Solution:

Let X be the waistline of a forty-year old male .

n= 36;

Sample mean of 36 male,

$$\Rightarrow \bar{X} \sim N\left(33, \frac{9}{36}\right) \text{ approximately by CLT.} \rightarrow \bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

(a) $P(\bar{X} > 31.5)$

$$= 1 - P(\bar{X} \le 31.5)$$

$$= 1 - P\left(Z \le \frac{31.5 - \mu_{\bar{X}}}{\sigma_{\bar{X}}}\right) \qquad z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$= 1 - P\left(Z \le \frac{31.5 - 33}{\sqrt{9/36}}\right) \rightarrow 3/\sqrt{36} \qquad = 1 - 0.0013$$

$$= 1 - P(Z \le -3.00) \qquad = 0.9987$$

Standard Normal Table

z	0.00	0.01	0.02
-3.1	0.0010	0.0009	0.0009
-3.0	0.0013	0.0013	0.0013
-2.9	0.0019	0.0018	0.0018

Continue…

(b) $P(32 \le \bar{X} \le 34)$

$$P(a \le \bar{X} \le b) = P(\bar{X} \le b) - P(\bar{X} \le a)$$

$$= P(\bar{X} \le 34) - P(\bar{X} \le 32)$$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$= P\left(Z \le \frac{34-33}{\sqrt{9/36}}\right) - P\left(Z \le \frac{32-33}{\sqrt{9/36}}\right)$$

$$= P(Z \le 2) - P(Z \le -2)$$

$$= 0.9772 - 0.0228$$

$$= 0.9544$$

Standard Normal Table

z	0.00	0.01	0.02
-2.2	0.0139	0.0136	0.0132
-2.1	0.0179	0.0174	0.0170
-2.0	0.0228	0.0222	0.0217

Standard Normal Table

z	0.00	0.01	0.02
1.8	0.9641	0.9649	0.9656
1.9	0.9713	0.9719	0.9726
2.0	0.9772	0.9778	0.9783

Continue…

33

(c) P(sample mean differs from population mean by more than 1)

$= P\left(\overline{X} - \mu < -1\right) + P\left(\overline{X} - \mu > 1\right)$ where μ $P(\bar{x} > 34) + P(\bar{x} < 32)$

$= P(\bar{X} < 32) + P(\bar{X} > 34)$

$P(\bar{x} < 32) + [1 - P(\bar{x} < 34)]$

$= 1 - P(32 \le \bar{X} \le 34)$

$= 1 - 0.9544$ from (b)

$= 0.0456$

Example 8.2-3

The body length (excluding the tail) of a particular species of mice is approximately normally distributed, with a mean of 12 cm and a standard deviation of 2.4 cm.

μ σ n

(a) If a random sample of 16 mice is selected, what is the probability that it will have an **average** body length of between 11 and 13 cm?

mean

$P(11 < \bar{x} < 13)$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Solution:

Let X be the body length of a mouse. $\therefore X \sim N(12, 2.4^2)$

(a) n= 16

Sample mean of 16 mice, $\bar{X} \sim N\left(12, \frac{2.4^2}{16}\right) \Rightarrow \bar{X} \sim N(12, 0.36)$ by CLT.

$$P(11 \le \bar{X} \le 13)$$
$$= P(\bar{X} \le 13) - P(\bar{X} \le 11)$$
$$= P\left(Z \le \frac{13-12}{\sqrt{0.36}}\right) - P\left(Z \le \frac{11-12}{\sqrt{0.36}}\right)$$
$$= P(Z \le 1.67) - P(Z \le -1.67)$$
$$= 0.9525 - 0.0475$$
$$= 0.9050$$

$\frac{\sigma}{\sqrt{n}} = \frac{2.4}{\sqrt{16}} = \sqrt{0.36} = 0.06$

Standard Normal Table

z	0.07	0.08	0.09
-1.8	0.0307	0.0301	0.0294
-1.7	0.0384	0.0375	0.0367
-1.6	0.0475	0.0465	0.0455

Standard Normal Table

z	0.07	0.08	0.09
1.4	0.9292	0.9306	0.9319
1.5	0.9418	0.9429	0.9441
1.6	0.9525	0.9535	0.9545

Continue…

(b) If a random sample of 25 mice is selected, what is the probability that it will have an **average** body length of between 11 and 13 cm?

(b) n= 25

Sample mean of 25 mice, $\bar{X} \sim N\left(12, \frac{2.4^2}{25}\right) \Rightarrow \bar{X} \sim N(12, 0.2304)$ by CLT

$$P(11 \le \bar{X} \le 13)$$

$$= P(\bar{X} \le 13) - P(\bar{X} \le 11)$$

$$= P\left(Z \le \frac{13-12}{\sqrt{0.2304}}\right) - P\left(Z \le \frac{11-12}{\sqrt{0.2304}}\right)$$

$$= P(Z \le 2.08) - P(Z \le -2.08)$$

$$= 0.9812 - 0.0188$$

$$= 0.9624$$

Standard Normal Table

z	0.07	0.08	0.09
-2.2	0.0116	0.0113	0.0110
-2.1	0.0150	0.0146	0.0143
-2.0	0.0192	0.0188	0.0183

Standard Normal Table

z	0.07	0.08	0.09
1.8	0.9693	0.9699	0.9706
1.9	0.9756	0.9761	0.9767
2.0	0.9808	0.9812	0.9817

Continue…

(c) Comment on the answers obtained in part (a) and (b).

The required probability becomes <u>larger</u> when sample size <u>increases</u>.