Date: 20 / 11 / 2023

لجان الدُفعات

Soul Academy

تفريغ إحصاء صيدلاني

موضوع المحاضرة : Lecture 14

رقم المحاضرة : فاينال 1

إعداد الصيدلانيه : نور مصطفى

The Sampling Distribution of a Sample Mean and the Central Limit Theorem

* الفكرة من هاد الشابتر انه نطلع (Zvalue)

* كما في السابق قلنا ...

population ← parameter

sample ← statics

Properties of the Sampling Distribution of $\bar{x}$

تقرأ (اكس بار)

- Let $X_1, X_2, \ldots\ldots, X_n$ be a random sample of size n drawn from any population with mean=μ and standard deviation = σ, then we have the following properties about the sampling distribution mean μ, that is:

خصائص μ ـة -

Properties of the Sampling Distribution of $\bar{x}$

1. Mean of the sampling distribution equals mean of sampled population*, that is,

true or false * ممكن يجي

$$\mu_{\bar{x}} = E(\bar{x}) = \mu.$$

sample لل

of population

2. Standard deviation of the sampling distribution equals

$$\frac{\text{Standard deviation of sampled population}}{\text{Square root of sample size}}$$

standard error of the mean

يطلق ب

of population

That is, $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

of sample

The standard deviation $\sigma_{\bar{x}}$ is often referred to as the **standard error of the mean**.

* كل ما standard error اقل معناها variation اقل •

Properties of the sampling distribution

1. The distribution of $\bar{x}$ will be normal.

   of sample

2. The mean $\mu_{\bar{x}}$ of the distribution of the values of μ $\bar{x}$ will be the same as the mean of the population from which the samples were drawn; $\boxed{\mu_{\bar{x}} = \mu}$.

3. The variance, $\sigma^2_{\bar{x}}$, of the distribution of $\bar{x}$ will be equal to the variance of the population divided by the sample size; $\boxed{\sigma^2_{\bar{x}} = \sigma^2 / n}$

$\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ اذا بدنا نطلع standard deviation بنحط الجذر وبتص

- The sample mean is a consistent estimator

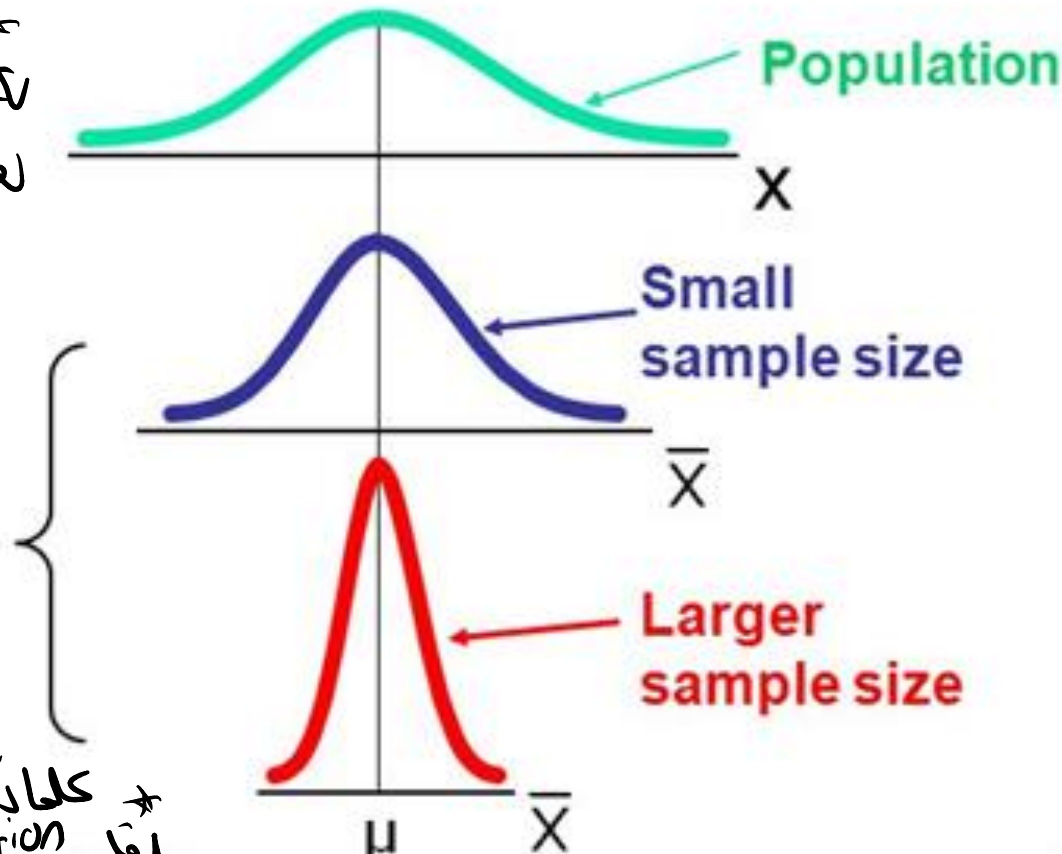
(the value of $\bar{x}$ becomes closer to μ as n increases):

* الـ٣ مخططات عندهم نفس mean
لكن value of $\bar{x}$ تكون أقرب للـ μ
لما n تزداد.

As n increases,

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

decreases



* كلما زدنا sample size كلما كان المنحنى أقل variation.

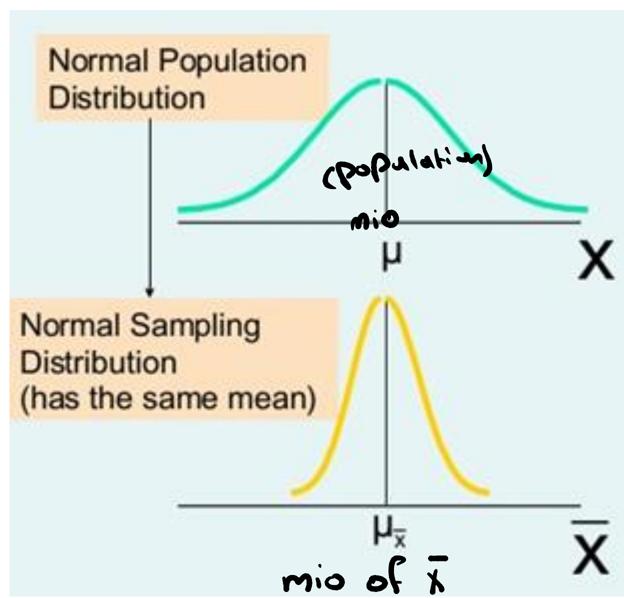
* عشان هيك اذا عملت تجربة و كان standard error عالي بنفضل نعيد التجربة ونزيد (n) عشان تقل error ...

Theorem

If a random sample of *n* observations is selected from a population with a normal distribution, the sampling distribution of $\bar{x}$ will be a normal distribution.

That is, if a probability distribution is normal, then the sampling distribution of the sample mean is exactly normal distribution, regardless of the sample size (n) small or large, with:

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



mio of $\bar{x}$
(samples)

→ بنعوضها مكان X
في معادلة z value

Sampling from Normal Populations

- **Central Tendency**

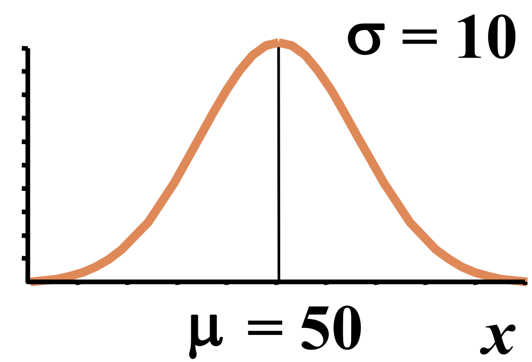
$$\mu_{\bar{x}} = \mu$$

- **Dispersion**

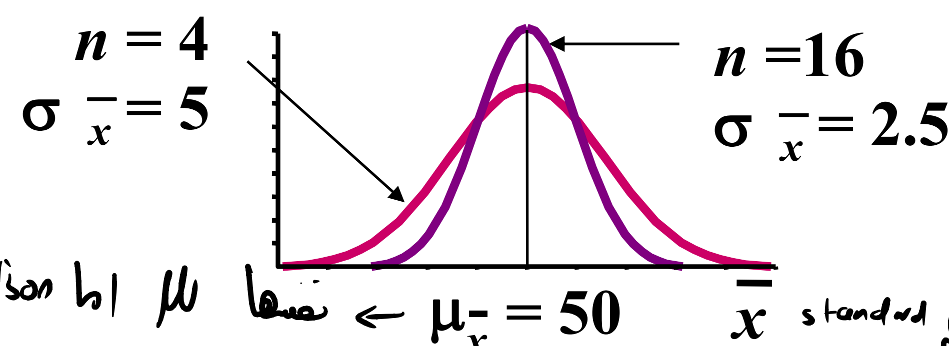
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

  – Sampling with replacement

Population Distribution



Sampling Distribution



يختلف despersion اما μ نفسها ←

standard error
أقل لأن n أكبر

Sampling from Normal Populations

- Note that:

If $n < 30$, then the sample size (n) is small.

If $n \geq 30$, then the sample size (n) is large.

If $X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$

$\boxed{\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)}$ for $n < 30$ or $n \geq 30$ # لانها normal

(مابتفرق سواء كانت اكبر من 30 او اصغر من 30)

Variance

* يعني بالمختصر لما يعطيني بالسؤال انه normal distribion ما بيهم بال sample size (n) ...

Sample distribution of the sample mean

Sampling from Normal Populations

- Example consider a normal population with $\mu = 50$ and $\sigma = 15$. suppose a random sample of size 9 is selected from this population, what is the sampling distribution for the sample mean?
- Solution: since the original population is normal, then the distribution of the sample mean is also (exactly) normal distribution with mean and variance as follows:

n

$\sigma_{\bar{x}}$

N(50,225)

$\mu_{\bar{x}} = \mu = 50$ because it's normal

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 15/3 = 5.$$

So for $\bar{X} \sim N(50, 25)$ is the sampling distribution of $\bar{X}$.

Sampling from Non-Normal Populations

- **Central Tendency**

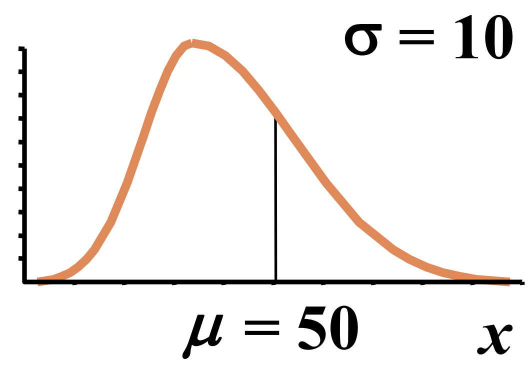
$$\mu_{\bar{x}} = \mu$$

- **Dispersion**

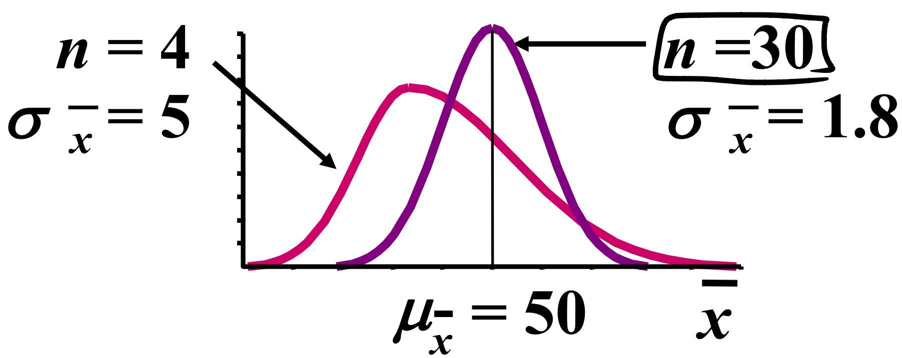
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

  - Sampling with replacement

Population Distribution



Sampling Distribution



Sample Mean Sampling Distribution: If the Population is **not** Normal

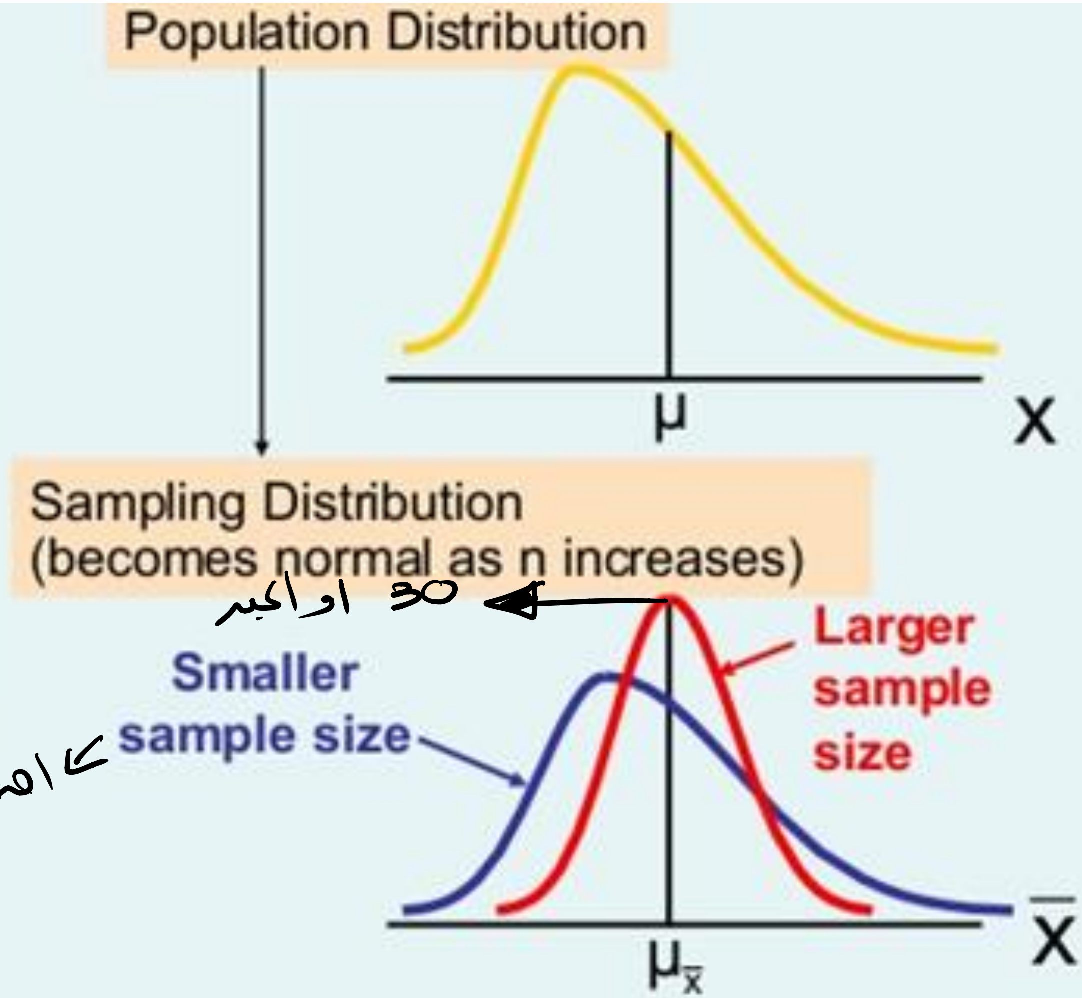
Sampling distribution properties:

Central Tendency

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

Variation

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



اصغر من 30 لحد مش normal

Central Limit Theorem (CLT)

Consider a random sample of *n* observations selected from a population (*any* probability distribution) with mean μ and standard deviation σ. Then, when *n* is sufficiently large, the sampling distribution of $\bar{x}$ will be approximately a normal distribution with mean $\mu_{\bar{x}} = \mu$ and standard deviation $\sigma_{\bar{x}} = \sigma / \sqrt{n}$.

أوريج الأفريج
لل sample

لل population

standard of error

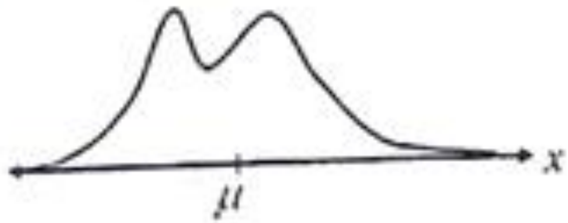
n

The larger the sample size, the better will be the normal approximation to the sampling distribution of $\bar{x}$

كل ما كانت sample size إلي أنا بختارها أكبر ، كل ما كان normal approximation to the sample distribution أفضل ...

large sample

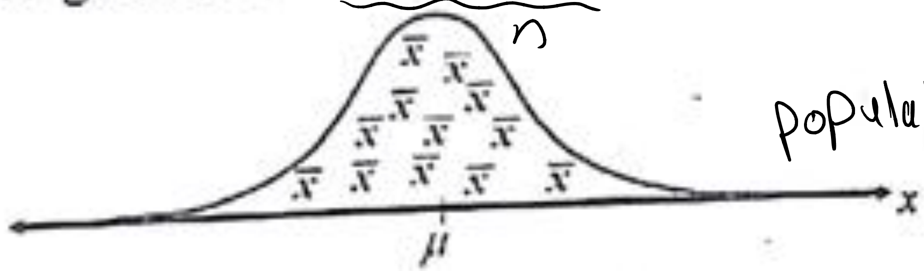
That is, if the sample size is large ($n \ge 30$), and the sample is drawn from any population with mean = μ and standard deviation = σ, as shown below:



then the sampling distribution of the sample mean approximates a normal distribution. The greater the sample size, the better the approximation, as shown below:

n

تصبح اقرب بال population



Any Population Distribution

Distribution of Sample Mean, $n \ge 30$

Central Limit Theorem (CLT)

- Example: Interpreting the CLT

فاتورة هاتف mean of sample هي نفسها μ

Phone bill for residents of a city have a **mean of $64** and a **standard deviation of $9**. random samples of **36 phone bills** are drawn from this population and the mean of each sample is determined. Find the mean and standard error of the mean of the sampling distribution. Then sketch a graph of the sampling of sample mean?

σ

Larg انه نعتبر 30 < (sample size) n لما ان #

normal distribution وانه

Solution: Interpreting the Central Limit Theorem

- The mean of the sampling distribution is equal to the population mean:

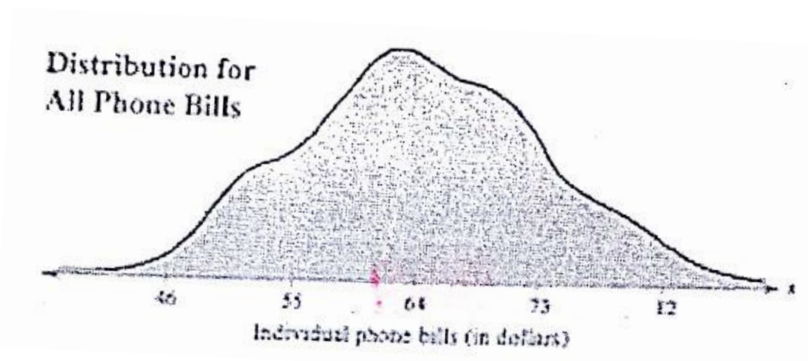
1 $$\mu_{\bar{x}} = \mu = 64$$

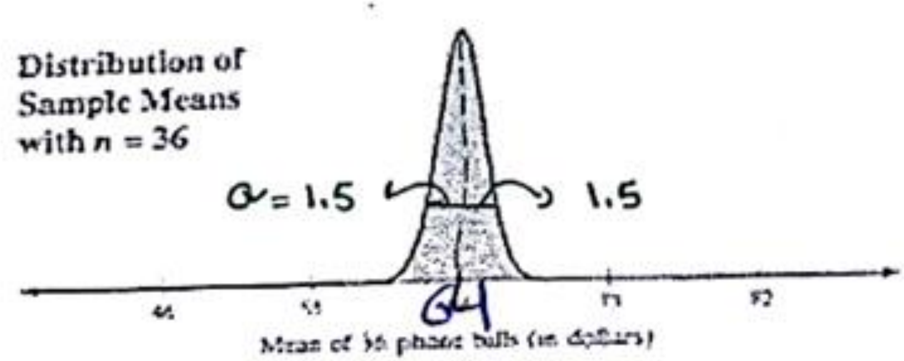
- The standard deviation (standard error of the mean) is equal to the population standard deviation divided by the square root of n:

2 $$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{36}} = 1.5$$

- Since the sample size is greater than 30, the sampling distribution can be approximated by a normal distribution with:

$$\mu_{\bar{x}} = 64 \qquad \sigma_{\bar{x}} = 1.5$$





*مثال ثاني نفس الفكرة ...

Central Limit Theorem (CLT)

- **Example**

هون موضحلك انه normal

The height of fully white oak are **normally distributed**, with a **mean of** μ **90** feet and **standard deviation of** σ **3.5** feet. Random samples of **size** n **4** are drawn from this population, and the mean of each sample is determined. Find the **mean** and **standard error** of the **mean of the sampling distribution**. Then sketch a graph of the sampling distribution of sample means?

نفس population

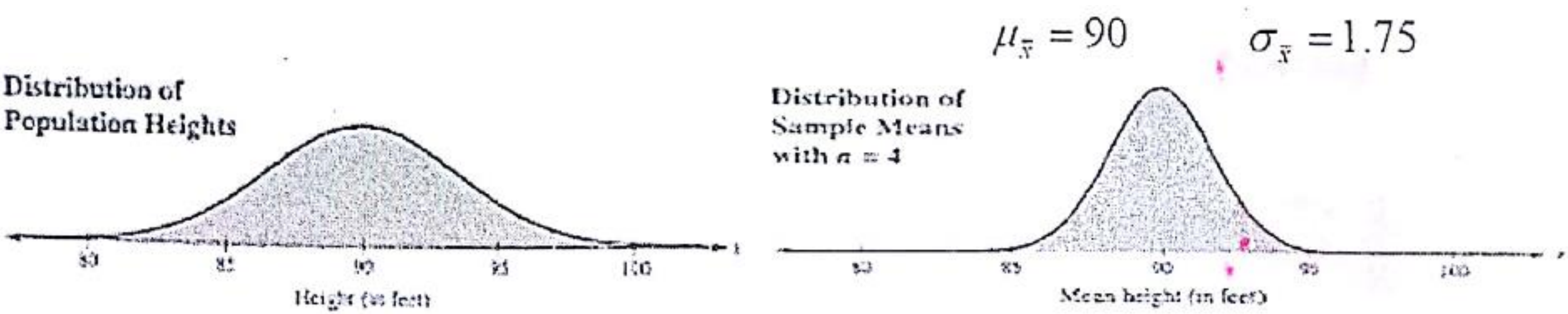
بس 50سلايد 😢

mean of population = mean of sample ← normal زي ما حكينا بال #
ولكن standard deviation بختلف ...

Solution

- The mean of the sampling distribution is equal to the population mean: $\mu_{\bar{x}} = \mu$ = 90
- The standard error of the mean is equal to the population standard deviation divided by root of n:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 3.5/2 = 1.75$$



The Probability of a Sample Mean

How to find the probability associated with a sample mean?

To find the probability for the sample mean, we transform X to a Z-score as follows:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$ is approximately $N(0, 1)$.

μ σ

mean

standard deviation

$\sigma_{\bar{X}}$

مش قلنا الفكرة من هاد السابق نطلع

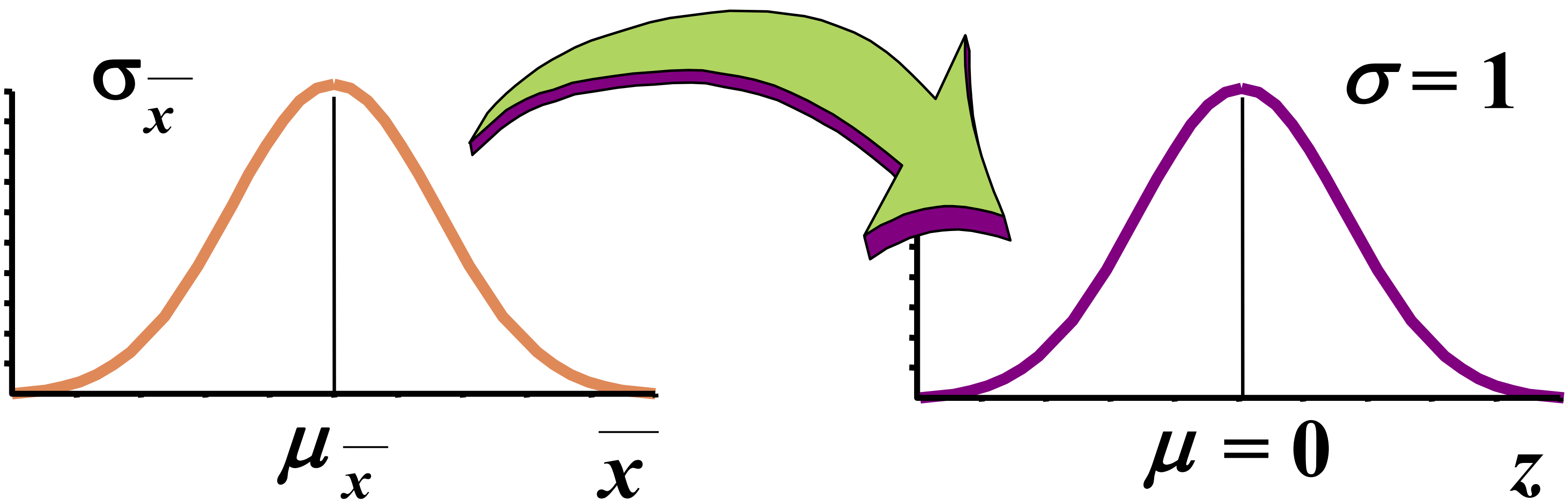
z value !!!

Standardizing the Sampling Distribution of $\bar{x}$

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

Sampling Distribution

Standardized Normal Distribution



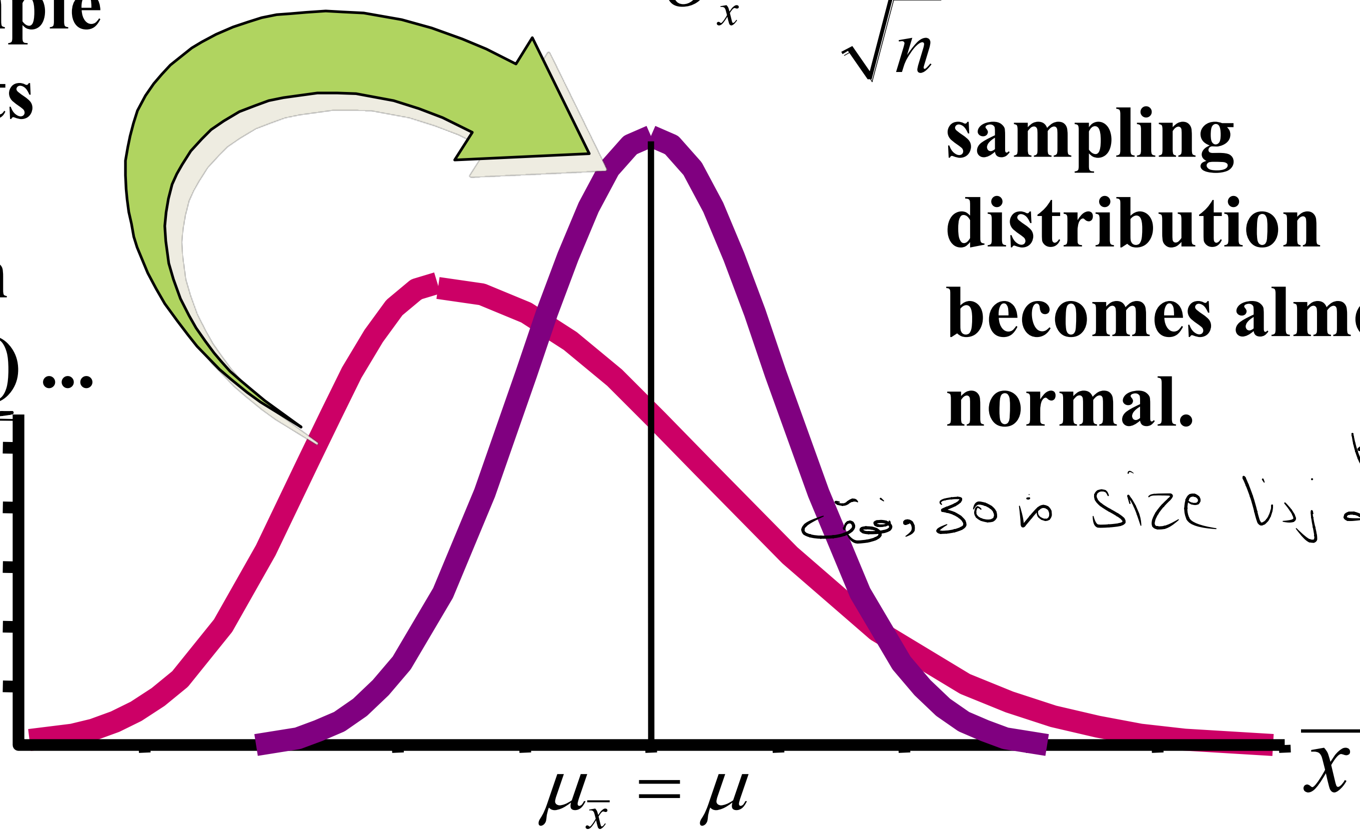
N(0,1)

Central Limit Theorem

As sample size gets large enough ($n \geq 30$) ...

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

sampling distribution becomes almost normal.



لما نزيد size من 30 وفوق

Examples

The graph shows the length of time people spend driving each day. You randomly select 50 drivers age 15 to 19. what is the probability that the mean time they spend driving each day is between 24.7 and 25.5 minutes?

Assume that $\mu = 25$ minutes and $\sigma = 1.5$ minutes?

Solution

From the central limit theorem (sample size is greater than 30), the sampling distribution of the sample mean is approximately normal with mean and standard deviation given as follows:

$$\mu_{\overline{x}} = \mu = 25 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 1.5/\sqrt{50} \approx 0.21213$$

اكبر 30
(normal)

Then the probability that the mean time they spend driving each day is between 24.7 and 25.5 minutes can be calculated as follows:

Normal Distribution

$\mu_{(xbar)} = 25 \quad \sigma_{(xbar)} = 0.21213$

Standard Normal Distribution

$\mu = 0 \quad \sigma = 1$

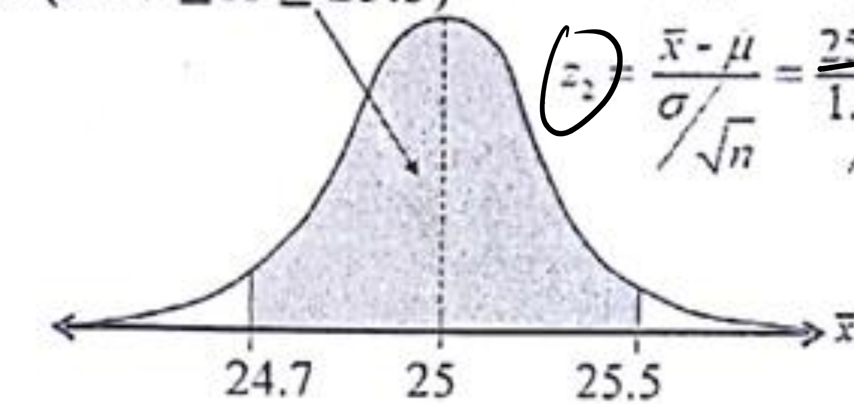
mean

$$z_1 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{24.7 - 25}{1.5/\sqrt{50}} = -1.41$$

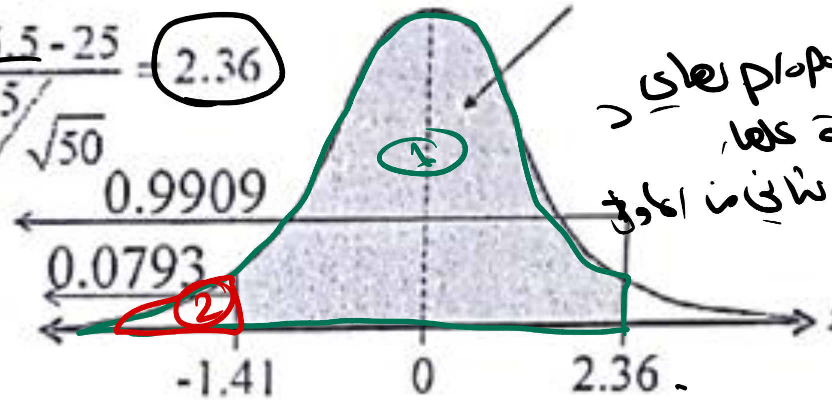
→0.21213

$$z_2 = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{25.5 - 25}{1.5/\sqrt{50}} = 2.36$$

$P(24.7 \le \overline{X} \le 25.5)$



$P(-1.41 \le Z \le 2.36)$



بدي propality هاي ل المنطقة كلها
مبطرح الثاني من الاول

بعطيك propality لليسار الرقم

$$P(24.7 \le \overline{X} \le 25.5) = P(-1.41 \le Z \le 2.36)$$
$$= N(2.36) - N(-1.41) = 0.9909 - 0.0793 = 0.9116$$

بنطلعها من جدول (مش معطى)

* نفس الفكرة

Examples

A bank auditor claims that credit card balances are normally distributed, with a mean of JD2870 and a standard deviation of JD900. you randomly select 25 credit card holders. What is the probability that their mean credit card balance is less than or equal to JD2500? أقل من 2500

Solution

Since the population is normally distributed, the sampling distribution of the sample mean is also normally distributed. Now, we are asked to find the probability associated with a sample mean and this can be calculated as follows:

$$\mu_{\bar{x}} = \mu = 2870$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{900}{\sqrt{25}} = 180$$

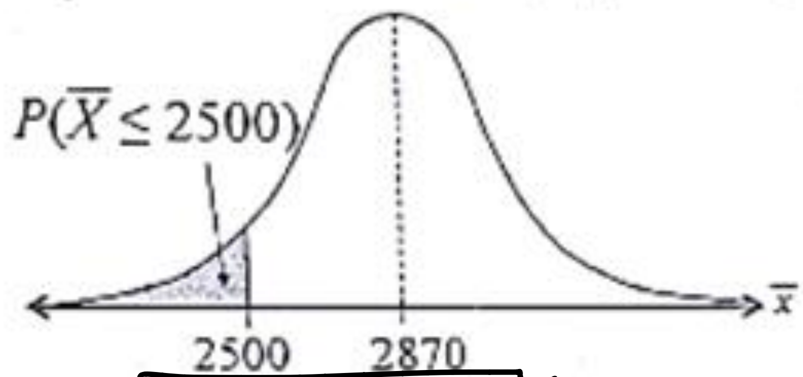
Normal Distribution

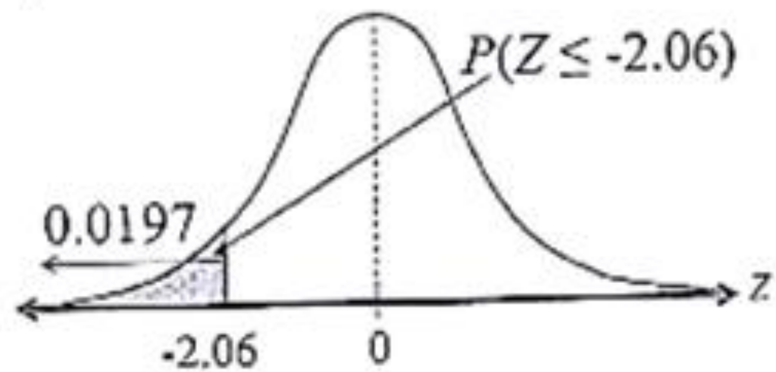
$\mu_{(xbar)} = 2870 \quad \sigma_{(xbar)} = 180$

Standard Normal Distribution

$\mu = 0 \quad \sigma = 1$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{2500 - 2870}{900 / \sqrt{25}} \approx -2.06$$





$P(\bar{X} \leq 2500) = P(Z \leq -2.06) = N(-2.06) = 0.0197 \rightarrow 0 - 0.0197 = 0.0197$

Conclusion: There is only a 2% chance that the mean of a sample of size 25 will have a balance less than or equal to JD2500 (*unusual event*). It is possible that the sample is unusual or it is possible that the auditor's claim that the mean is JD2870 is incorrect.

Examples

Suppose that the TOFEL exam scores for pharmacy students graduated from the Hashemite University are normally distributed with $\mu = 500$ and $\sigma = 100$.

ما يهمنا انه n اقل من 30 لأنه
حكالي بالسؤال انه normal distribsion

In a random sample of size $n = 25$ students, what is the probability that the sample mean would be greater than 540? $\rightarrow \bar{x} > 540$

(المساحة إلى عاليمين)

Solution

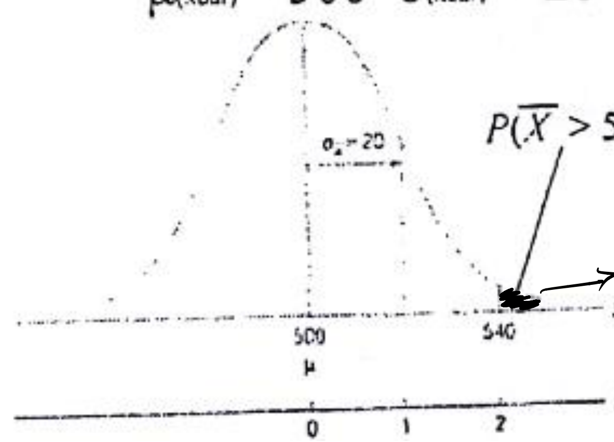
Since the population is normally distributed, the sampling distribution of the sample mean is also normally distributed. Now, we are asked to find the probability associated with a sample mean and this can be calculated as follows:

Normal Distribution

$\mu_{(xbar)} = 500 \quad \sigma_{(xbar)} = 20 \rightarrow \sigma_{\bar{x}} = \frac{100}{\sqrt{25}}$

$\sigma_{\bar{x}} = 20$

$$\begin{aligned}
P(\bar{X} > 540) &= 1 - P(\bar{X} \le 540) \\
&= 1 - P(Z \le (\underset{\bar{x}}{540} - \underset{\mu}{500})/\underset{\sigma_{\bar{x}}}{20}) \\
&= 1 - P(Z \le 2) \\
&= 1 - N(2) \\
&= 1 - 0.9772 \\
&= 0.0228
\end{aligned}$$



Thinking Challenge

You're an operations analyst for AT&T. Long-distance telephone calls are normally distributed with $\boxed{\mu = 8}$ min. and $\boxed{\sigma = 2}$ min. If you select random samples of $\boxed{25}$ calls, what percentage of the **sample means** would be between **7.8** & **8.2** minutes?

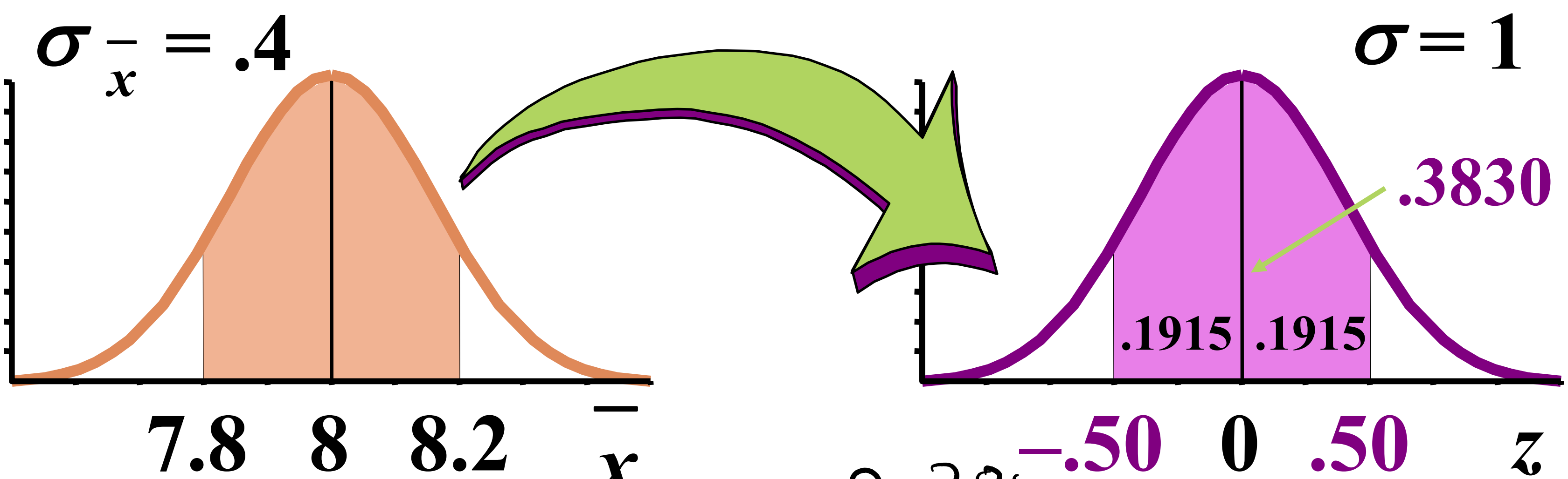
Sampling Distribution Solution*

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{7.8 - 8}{2 / \sqrt{25}} = -.50$$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{8.2 - 8}{2 / \sqrt{25}} = .50$$

Sampling Distribution

Standardized Normal Distribution

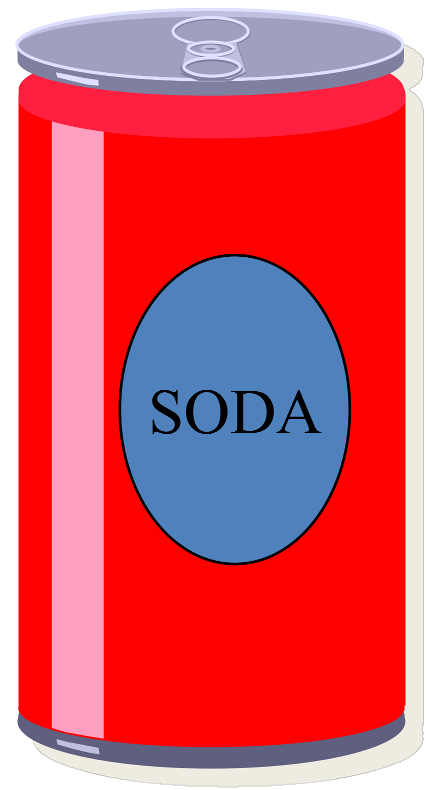


الجواب ← 38.3%

Central Limit Theorem Example

The amount of soda in cans of a particular brand has a $\overset{\mu}{\text{mean}}$ of **12** oz and a standard deviation of $\overset{\sigma}{\mathbf{.2}}$ oz. If you select random samples of $\overset{n}{\mathbf{50}}$ cans, what percentage of the **sample means** would be less than **11.95** oz?

< 11.95

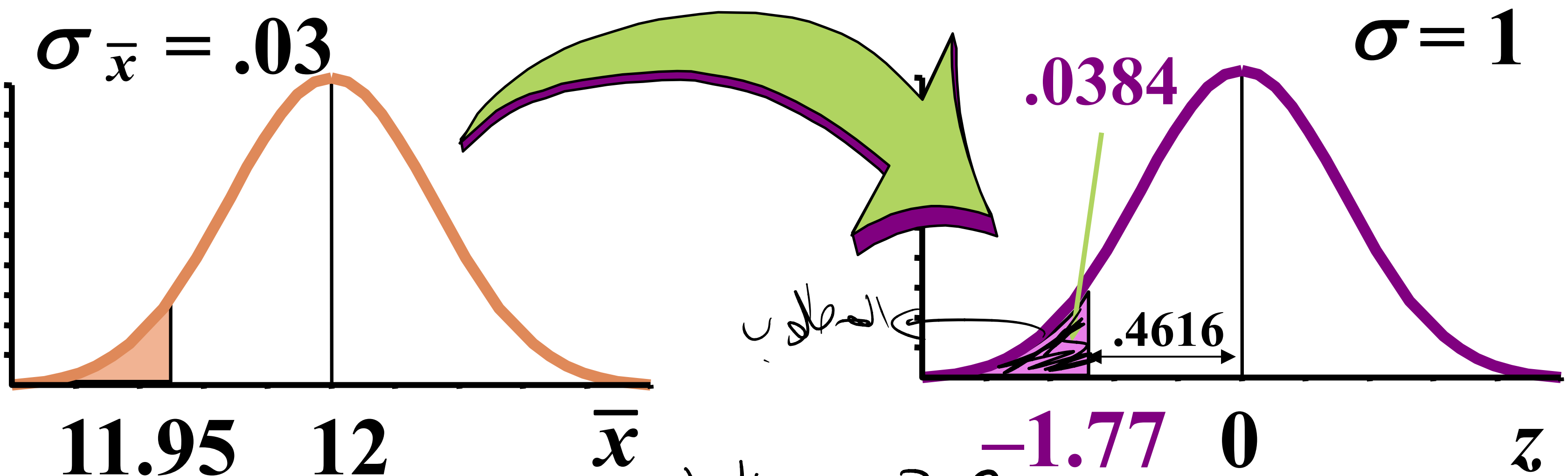


Central Limit Theorem Solution*

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{11.95 - 12}{.2 / \sqrt{50}} = -1.77$$

Sampling Distribution

Standardized Normal Distribution



الجواب 3.84%

Shaded area exaggerated

Sampling Distributions

* هذا السلايد revision ...

Parameter & Statistic

A **parameter** is a numerical descriptive measure of a population. Because it is based on all the observations in the population, its value is almost always unknown.

A **sample statistic** is a numerical descriptive measure of a sample. It is calculated from the observations in the sample.

Common Statistics & Parameters

هذا الجدول للحسابات . . .

	Sample Statistic	Population Parameter
Mean	$\overline{x}$	μ
Standard Deviation	s	σ سيغما
Variance	s^2	σ^2
Binomial Proportion	p-hat $\hat{p}$	p

The Concept of Sample Distribution

- In sample distributions, we are moving from descriptive statistics to inferential statistics. Inferential statistics allow the researcher to come to conclusion about a populations on the basis of descriptive statistics about a sample.
- In real life parameters of populations are known and unknowable.
- Rather than investigating the whole population, we take a sample, calculate a statistics related to the parameter of interest, and make an inference.
- The sampling distribution of the statistics tells us how the value of the statistic varies from one sample to another.

Sampling Distribution

The **sampling distribution** of a sample statistic calculated from a sample of *n* measurements is the probability distribution of the statistic in all possible samples of size n taken from the same population.

The **sampling distribution** are used to calculate the probability that sample statistics could have occurred by chance and thus to decide whether something that is true of sample statistics is also likely to be true of a population parameter.

Developing
Sampling Distributions

افترض

Suppose There's a Population ...

- Population size, $N = 4$
- Random variable, x
- Values of x: 1, 2, 3, 4
- Uniform distribution

Population Characteristics

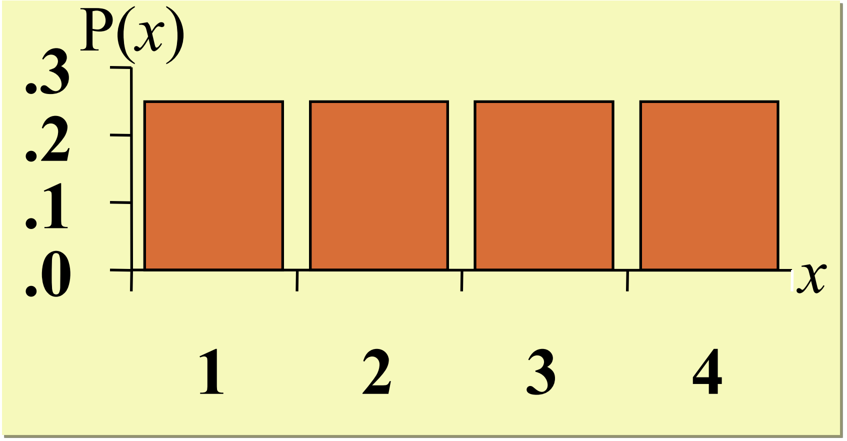
Summary Measure

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = 2.5$$

$$\mu = \frac{1+2+3+4}{4} = 2.5$$

2.5 or 25% لكل وحدة propulity ال

Population Distribution



All Possible Samples of Size $n = 2$

16 Samples

1st Obs	2nd Observation 1	2	3	4
1	1,1	1,2	1,3	1,4
2	2,1	2,2	2,3	2,4
3	3,1	3,2	3,3	3,4
4	4,1	4,2	4,3	4,4

16 Sample Means

1st Obs	2nd Observation 1	2	3	4
1	1.0	1.5	2.0	2.5
2	1.5	2.0	2.5	3.0
3	2.0	2.5	3.0	3.5
4	2.5	3.0	3.5	4.0

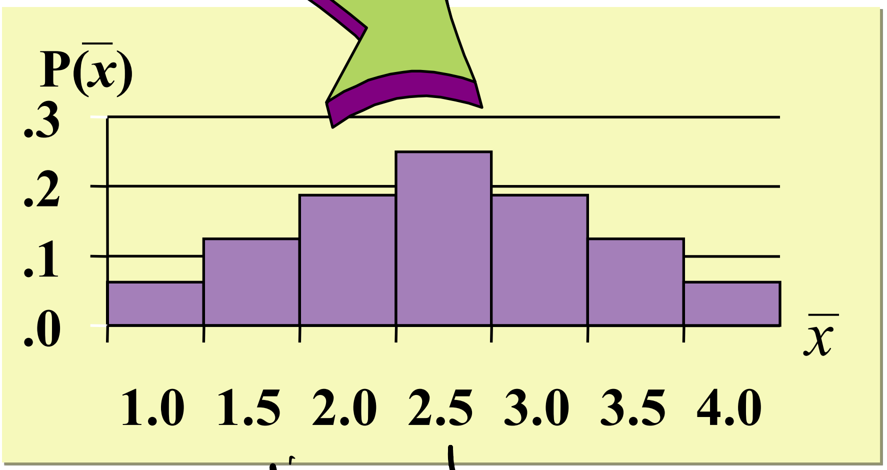
Sample with replacement

Sampling Distribution of All Sample Means

16 Sample Means

1st Obs	2nd Observation			
	1	2	3	4
1	1.0	1.5	2.0	2.5
2	1.5	2.0	2.5	3.0
3	2.0	2.5	3.0	3.5
4	2.5	3.0	3.5	4.0

Sampling Distribution of the Sample Mean



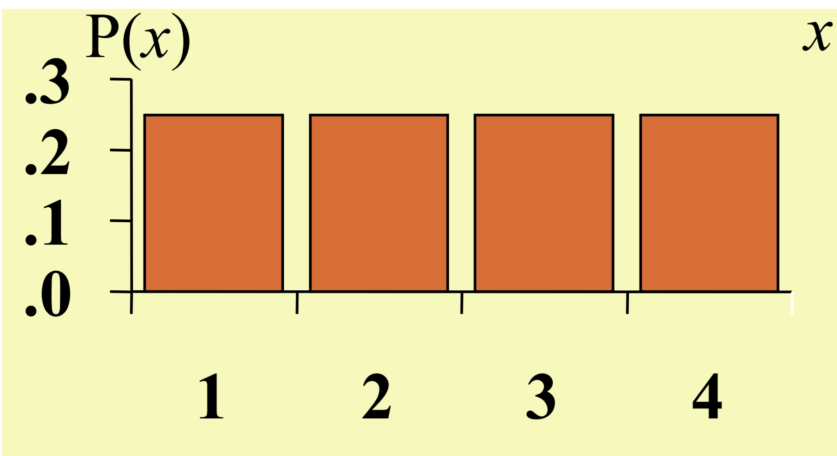
median

Summary Measure of All Sample Means

$$\mu_{\bar{X}} = \frac{\sum_{i=1}^{N} \bar{x}_i}{N} = \frac{1.0 + 1.5 + \ldots + 4.0}{16} = 2.5$$

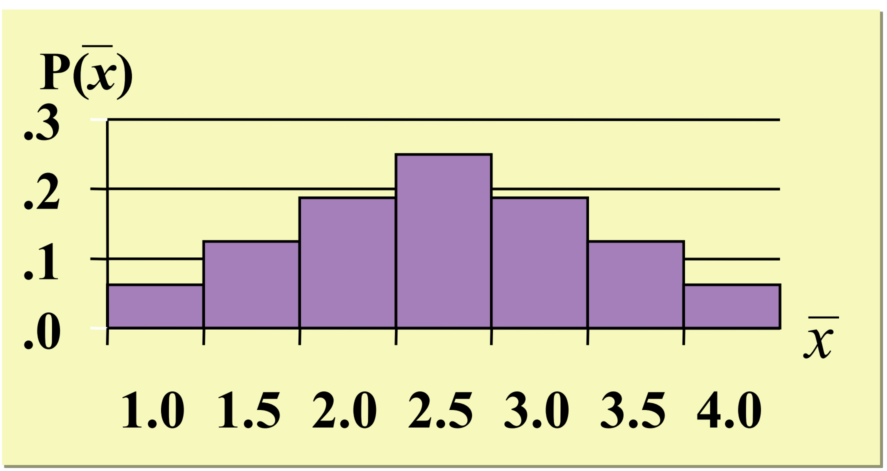
Comparison

Population



$$\mu = 2.5$$

Sampling Distribution

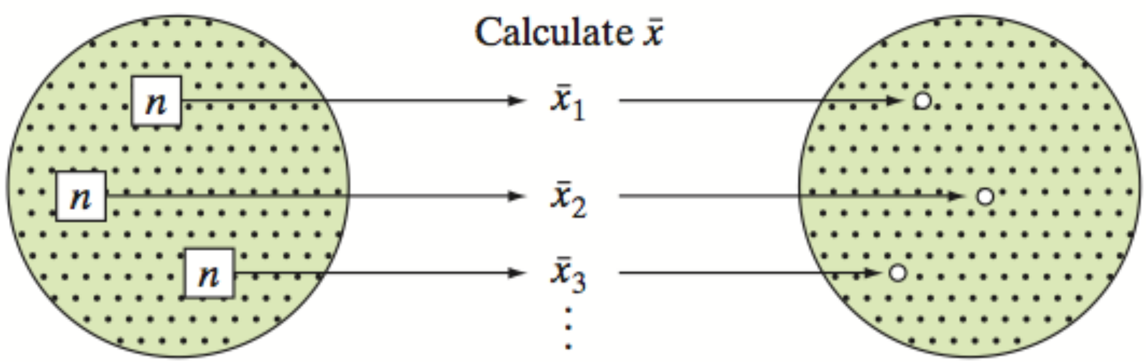


$$\mu_{\bar{x}} = 2.5$$

Key Ideas

Generating the Sampling Distribution of $\bar{x}$

Select sample size n (large)
from target population



Population:
Mean = μ
Std. Dev. = σ
Unknown shape

Repeat this process an infinite number of times

Sampling distribution of $\bar{x}$
(i.e., theoretical population of $\bar{x}$'s)
Mean = $\mu_{\bar{x}} = \mu$
Std. Dev. = $\sigma_{\bar{x}} = \sigma/\sqrt{n}$
Normal distribution (Central Limit Theorem)

← اسم النظرية التي تعتمد على
normal distribution

The Sampling Distribution of the Sample Proportion $\hat{p}$

of population p

هانة يابطل ...

Introduction

When conducting research about a population, researchers in business, socioeconomics, politics, medicine, nursing and pharmaceutical matters are often more interested in the proportion of a population (p) with a particular characteristic, rather than the number of population elements with the characteristic, for example:

- Proportion of population who support the policies and economic plans for the government in Jordan.
- Proportion of students who support the 4 days study university system in Hashemite University.
- Proportion of manufactured objects that are defect free.
- Proportion of employees with extended health care plans.
- Percentage of the labour force that is unemployed.
- Proportion of students suffering from the flu.

In each of these situation, the actual number of population elements with the characteristic will vary with the sample size. But the aim of obtaining samples is to estimate the proportion (or percentage) of the population with the characteristic of interest.

The Sampling Distribution of the Sample Proportion $\hat{p}$

نتمنى

- We hope that the sample proportion is close to the population proportion.
- How close can we expect it to be?
- Would it be worth it to collect a larger sample?
  - If the sample were larger, would we expect the sample proportion to be closer to the population proportion?
  - How much closer?

How to Calculate the sample Proportion ($p^{\wedge}$)

Suppose that p is the proportion of a population with a particular characteristic, then if p is unknown, we can estimate it by using the sample proportion $p^{\wedge}$ (p-hat) using the following steps:

1. Draw a random sample (SRS) of size n elements from the population that contains N elements.
2. Let X be the number of sample elements with the characteristics (number of successes).
3. Calculate the sample proportion ($p^{\wedge}$) using the following formula:

$$p^{\wedge} = \frac{count\ of\ successes\ in\ the\ sample}{n} = \frac{X}{n}$$

That is, $p^{\wedge}$ is the proportion of elements of the sample of size n that have the characteristic of interest

Example

The Jordanian Ministry of Health did a survey of in a random sample of 10904 Jordanian peoples and smoking habits. The researchers defined :

Frequent smoking as having 5 or more cigarettes in a row three or more times in the past two hours. According to this definition, 2486 persons were classified as frequent smoking. Based on these data, estimate the proportion p of all Jordanian peoples who admits to frequent smoking?

Solution

The true proportion (p) of Jordanian peoples who are frequent smoking can be estimated by using the sample proportion ($\hat{p}$) which can be calculated as follows:

$$\hat{p} = \frac{\textit{Number of Jordanian persons who are classified as frequent smokin}}{\textit{Number of Jordanian persons in the sample}}$$

$$\Longrightarrow \quad \hat{p} = \frac{X}{n} = \frac{2486}{10904} = 0.228$$

The Sampling Distribution of the Sample Proportion $p^\wedge$

- Choose an SRS of size n from a large population that contains population proportion p of successes. Let $p^\wedge$ be the sample proportion of successes,

$$p^\wedge = \frac{\text{count of successes in the sample}}{n}$$

- Then :

➢ As the sample size increases, the sampling distribution of p becomes approximately normal.

➢ The **mean** of the sampling distribution is **p**.

➢ The standard deviation of the sampling distribution is

$$\boxed{\text{Standard deviation of } \hat{p} = \sqrt{\frac{p(1-p)}{n}}} \to \varphi$$

* لو طلب بالسؤال نحسب propility نحول ل zvalue .

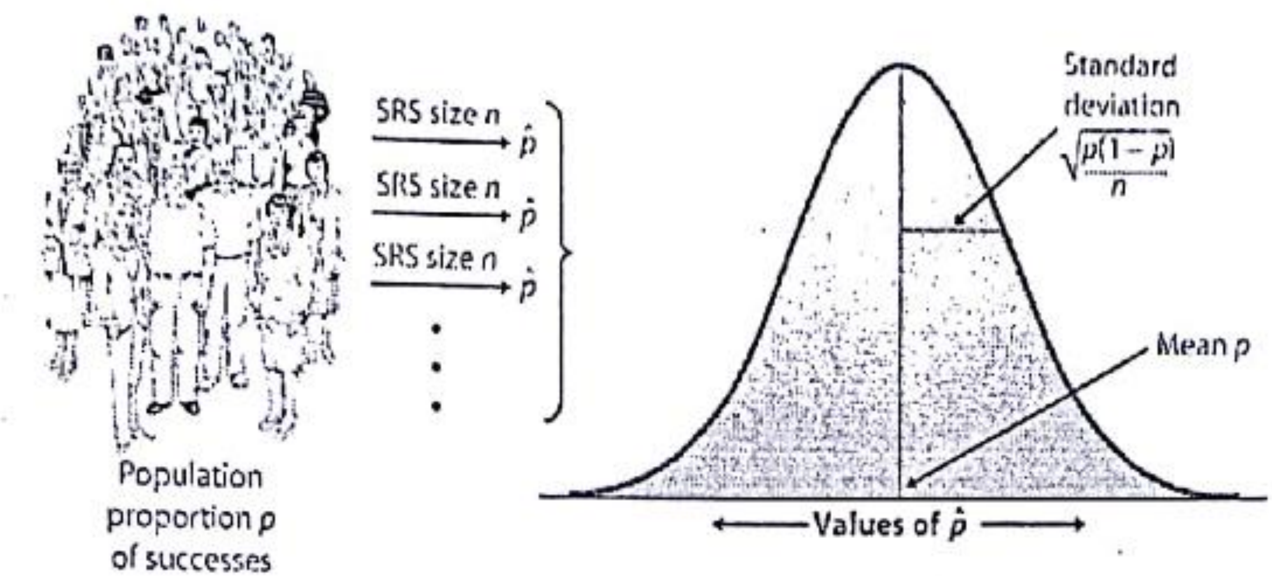
The Sampling Distribution of *p*^

- It turns out that the sampling distribution of *p*^ is approximately normal with the following parameters.

$$\text{Mean of } \hat{p} = p \quad \rightarrow \mu_{\hat{p}} = p$$

$$\text{Variance of } \hat{p} = \frac{p(1-p)}{n}$$

$$\text{Standard deviation of } \hat{p} = \sqrt{\frac{p(1-p)}{n}}$$



The Sample Proportion

- Let *p* be the population proportion.
- Then *p* is a fixed value (for a given population).
- Let $\hat{p}$ ("*p*-hat") be the sample proportion.
- Then $\hat{p}$ is a random variable; it takes on a new value every time a sample is collected.
- The sampling distribution of $\hat{p}$ is the probability distribution of all the possible values of $\hat{p}$.

normal

The Central Limit Theorem for the Sample Proportion ($\hat{p}$)

For any population proportion (p), the sampling distribution of the sample proportion ($\hat{p}$) is *approximately normal if the sample size (n) is sufficiently large*. As a general guideline, the normal distribution approximation is justified when $np \geq 5$ and $n(1 - p) \geq 5$, that is:

Sampling Distribution of the Sample Proportion

$$\hat{p} \ \widetilde{Approximately} \ \boxed{N\left(p\,,\frac{p(1-p)}{n}\right)} \text{ if } n \text{ is sufficiently large.}$$

- The approximation to the normal distribution is excellent if:

$$\boxed{np \geq 5 \text{ and } n(1-p) \geq 5.}$$

لـ اذا اختل اي شرط منهم مابقدر اعمل approximation to the normal distribution.

Example

Suppose that the population of interest is the Hashemite University pharmacy students. Assume that the proportion of students in the population mean of $\hat{p}$ who wear eyeglasses is **p=0.25**. if a random sample (SRS) of **50** students is to be selected, then define the characteristics of the sampling distribution of the sample proportion ***p*^**, where ***p*^** is the proportion of HU pharmacy students in the random sample who wear eyeglasses?

Example

Solution

We know that the mean of the sampling distribution of $\hat{p}$ is:

$$\mu_{\hat{p}} = p = 0.25$$

and we know that the variance and standard deviation of the sampling distribution of $\hat{p}$ are:

$$\sigma^2_{\hat{p}} = \frac{p(1-p)}{n} = \frac{(0.25)(0.75)}{50} = 0.00375 \quad ; \quad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{0.00375} = 0.0612$$

n p اكبر من 5

Finally, for sample size of $n = 50$, we have $np = (50)(0.25) = 12.5$ and

اكبر من 5

$n(1-p) = (50)(0.75) = 37.5$ thus the large sample conditions are satisfied and the sampling distribution of $\hat{p}$ can be approximated by a normal distribution, that is: $\hat{p}$ *Approximately* $N\left(p\,,\frac{p(1-p)}{n}\right) = N(0.25\,, 0.00375)$

How to Find the Probability associated with a Sample Proportion ($\hat{p}$)?

Since the sampling distribution of $\hat{p}$ can be approximated by a normal distribution, we can use the areas under the curve of the standard normal distribution to answer the probability question. With the normal distribution, we need to compute a Z-score value, where

$$Z = \frac{\hat{p} - \mu_{\hat{p}}}{\sigma_{\hat{p}}} = \frac{\hat{p} - P}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

Notation: The above Z-score is identical to the Z value we have always used, with the exception of the notation $\hat{p}$ and $\sigma_{\hat{p}}$, which is used to remind us we are dealing with the sampling distribution of the sample proportion ($\hat{p}$).

مدر للبول

Example

Thiazide diuretics are often the first, but not the only choice in high blood pressure medications. Suppose that

mean   يفضلون

20% of all doctors in Jordan favour Thiazide diuretics. If the manger of a pharmaceutical industries company in Jordan take a random sample (SRS) of size

sample

n = 600 doctors, then what is the $\boxed{\text{probability}}$ that the sample proportion (***p***^) of doctors who favour Thiazide diuretics will be:

1. Between 0.18 and 0.22? $P(0.18 \leq \hat{p} \leq 0.22)$
2. Less than or equal to 0.18? $P(\hat{p} \leq 0.18)$
3. More than 0.22? $P(\hat{p} \geq 0.22)$

Solution

We know that the mean of the sampling distribution of $\hat{p}$ is:

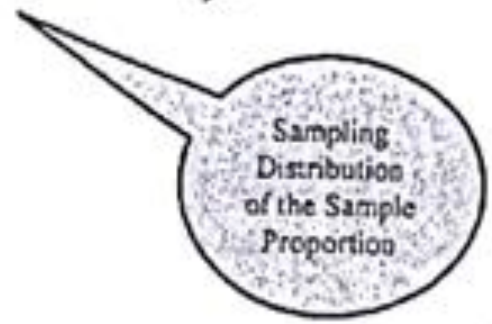
$$\mu_{\hat{p}} = p = 20\% = 0.2$$

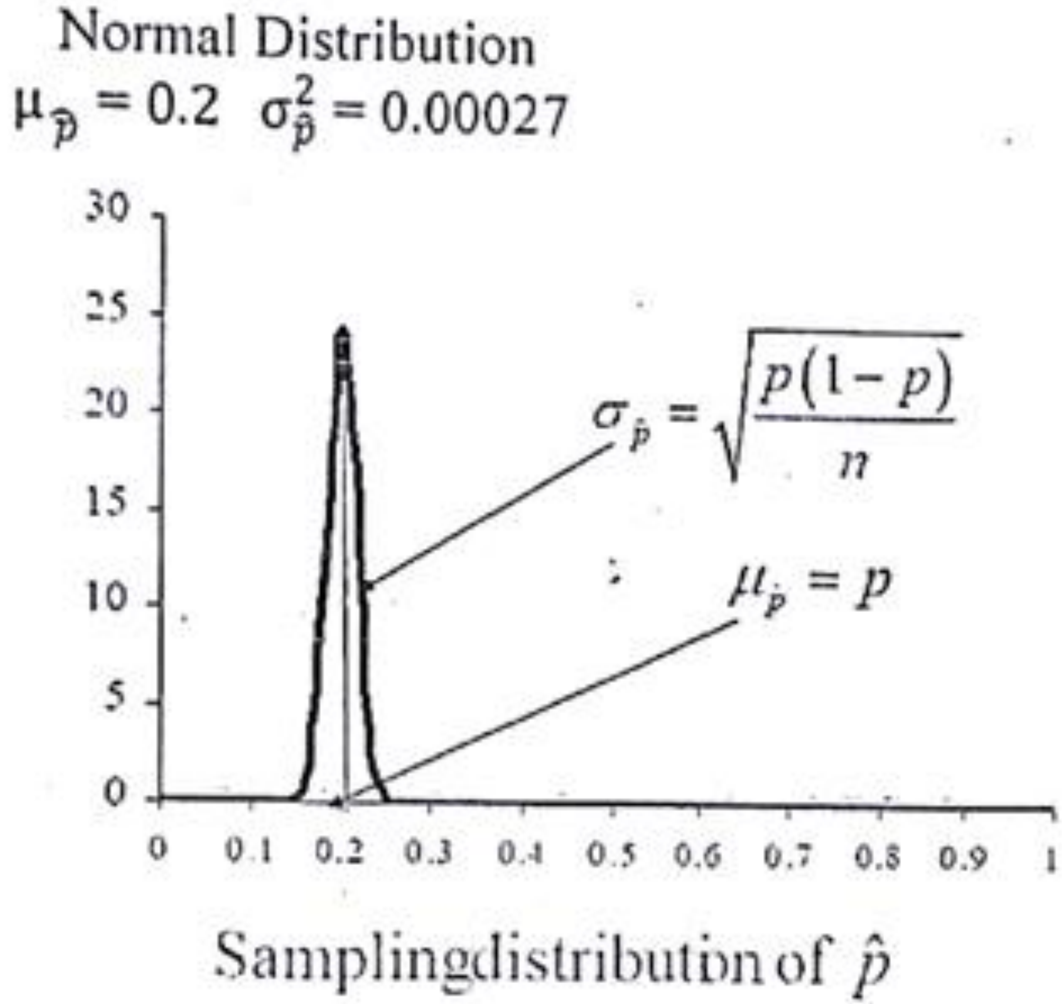
and we know that the variance and standard deviation of the sampling distribution of $\hat{p}$ are:

$$\sigma^2_{\hat{p}} = \frac{p(1-p)}{n} = \frac{(0.2(0.8)}{600} = 0.00027 \quad ; \quad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{0.00027} = 0.0164$$

اكيمن 5

Finally, for sample size of $n = 600$, we have $np = (600)(0.2) = 120$ and اكيمن 5 $n(1-p) = (600)(0.8) = 480$; thus the large sample conditions are satisfied and the sampling distribution of $\hat{p}$ can be approximated by a normal distribution, that is: $\hat{p}$ $\widetilde{Approximately}$ $N\left(p\,,\frac{p(1-p)}{n}\right) = N(0.2\,, 0.00027)$.





Now, the required probabilities can be found as follows:

(i) Probability that the sample proportion ($\hat{p}$) of doctors who favor Thiazide diuretics will be between 0.18 and 0.22 can be calculated as follows:

$$P(0.18 \le \hat{p} \le 0.22) = P\left(\frac{0.18 - 0.2}{0.0164} \le \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \le \frac{0.22 - 0.2}{0.0164}\right)$$

$$= P(-1.22 \le Z \le 1.22)$$

$$= N(1.22) - N(-1.22) = 0.8888 - 0.1112$$

$$= 0.7776$$

(ii) Probability that the sample proportion ($\hat{p}$) of doctors who favor Thiazide diuretics will be less than or eq 18 can be calculated as follows:

$$P(\hat{p} \le 0.18) = P\left(\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \le \frac{0.18 -}{0.0164}\right.$$

$$= P(Z \le -1.22) = N(-1.22) = 0.1112$$ $\rightarrow 0 - 0.1112 = 0.1112$

(iii) Probability that the sample proportion ($\hat{p}$) of doctors who favor Thiazide diuretics will be more than 0.22 can be calculated as follows:

$$P(\hat{p} > 0.22) = 1 - P(\hat{p} \le 0.22) = 1 - P\left(\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \le \frac{0.22 - 0.2}{0.0164}\right)$$

$$= 1 - P(Z \le 1.22) = 1 - N(1.22)$$

$$= 1 - 0.8888 = 0.1112$$