# Statistics

# Lecture 10

Feb 19-8:47 AM

SG 19

## Central – Limit Theorem CLT

$\mu_{\bar{x}} = \mu$ $\quad \sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$ $\quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

Clear all lists. [2nd] [+] [4:ClearAllLists] [Enter]

Store 2,4,6,8 in L1.

Use [1-Var Stats] with L1 only to find

$\mu = 5$ ✓ $\quad \sigma = 2.236$ $\quad \sigma^2 = 5$ ✓

[VARS] [5:Statistics] [4:$\sigma_x$] [$x^2$] [Enter]

Take all Samples of [size 2] with replacement from this data.

| | | | |
|---|---|---|---|
| 2,2 | 2,4 | 2,6 | 2,8 |
| 4,2 | 4,4 | 4,6 | 4,8 |
| 6,2 | 6,4 | 6,6 | 6,8 |
| 8,2 | 8,4 | 8,6 | 8,8 |

Find $\bar{x}$ of these Samples

| | | | |
|---|---|---|---|
| 2 | 3 | 4 | 5 |
| 3 | 4 | 5 | 6 |
| 4 | 5 | 6 | 7 |
| 5 | 6 | 7 | 8 |

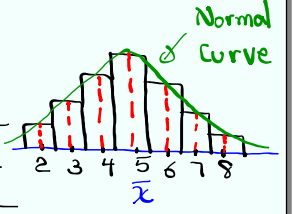
May 2-8:08 AM

| | | | |
|---|---|---|---|
| 2 | 3 | 4 | 5 |
| 3 | 4 | 5 | 6 |
| 4 | 5 | 6 | 7 |
| 5 | 6 | 7 | 8 |

16 means

| $\bar{x}$ | $P(\bar{x})$ |
|---|---|
| 2 | 1/16 |
| 3 | 2/16 |
| 4 | 3/16 |
| 5 | 4/16 |
| 6 | 3/16 |
| 7 | 2/16 |
| 8 | 1/16 |

Prob. dist. histogram




$\bar{x} \rightarrow L2$, $P(\bar{x}) \rightarrow L3$

use [1-Var Stats] with L2 & L3 to find

$\mu = 5$ $\sigma = 1.581$ $\sigma^2 = 2.5 = \frac{5}{2}$

**CLT**

$\mu_{\bar{x}} = \mu = 5$

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{5}{2}$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.236}{\sqrt{2}} = 1.581$

May 2-8:18 AM

Clear all lists

Store 2, 4, 6, 8, 10 in L1.

use [1-Var Stats] with L1 only to find

[$\mu = 6$] ✓ $\sigma = 2.828$ [$\sigma^2 = 8$] ✓

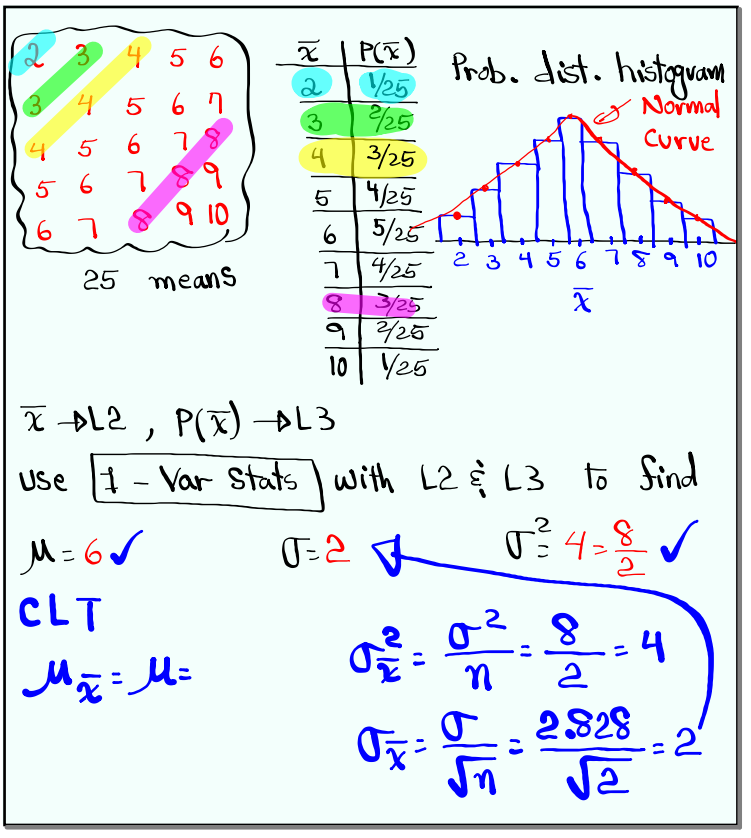
Let's take all Samples of Size 2 with replacement from this group.

| | | | | |
|---|---|---|---|---|
| 2,2 | 2,4 | 2,6 | 2,8 | 2,10 |
| 4,2 | 4,4 | 4,6 | 4,8 | 4,10 |
| 6,2 | 6,4 | 6,6 | 6,8 | 6,10 |
| 8,2 | 8,4 | 8,6 | 8,8 | 8,10 |
| 10,2 | 10,4 | 10,6 | 10,8 | 10,10 |

Find $\bar{x}$ of each Sample.

| | | | | |
|---|---|---|---|---|
| 2 | 3 | 4 | 5 | 6 |
| 3 | 4 | 5 | 6 | 7 |
| 4 | 5 | 6 | 7 | 8 |
| 5 | 6 | 7 | 8 | 9 |
| 6 | 7 | 8 | 9 | 10 |

May 2-8:18 AM




| $\bar{x}$ | $P(\bar{x})$ |
|---|---|
| 2 | 1/25 |
| 3 | 2/25 |
| 4 | 3/25 |
| 5 | 4/25 |
| 6 | 5/25 |
| 7 | 4/25 |
| 8 | 3/25 |
| 9 | 2/25 |
| 10 | 1/25 |

May 2-8:36 AM

Consider a normal Prob. dist with $\mu = 88$ and $\sigma = 12$.

Suppose we take all Samples of Size 4.

$\mu_{\bar{x}} = \mu = 88$

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{4}} = \frac{12}{2} = 6$

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{12^2}{4} = 36$

CLT ✓

May 2-8:45 AM

Salaries of nurses are normally distributed with the mean of \$7400/mo. and standard deviation of \$600/mo.

If we take all samples of 16 nurses,

1) find $\mu_{\bar{x}} = \mu = \boxed{7400}$ CLT

2) find $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{600}{\sqrt{16}} = \frac{600}{4} = \boxed{150}$

May 2-8:49 AM

Scores on a math exam are normally dist. with $\mu = 85$ and $\sigma = 10$.

If we randomly select $n=4$ 4 exams, find the prob. that $\bar{x}$ their mean Score is between 75 and 90.

$P(75 < \bar{x} < 90)$


= normalcdf(75,90,85,5)

= $\boxed{.819}$ = 81.9% ≈ 82%

CLT $\begin{cases} \mu_{\bar{x}} = \mu = 85 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{4}} = 5 \end{cases}$

May 2-8:53 AM

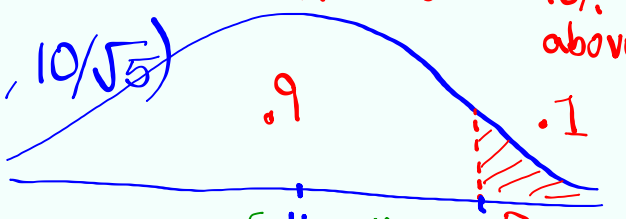
Find $\bar{x} = P_{90}$ for randomly selected 5 exams.

Round to whole #.

$\bar{x} = \text{invNorm}(.9, 85, 10/\sqrt{5})$

$= 90.731$

$\approx 91$




SG 19

CLT $\begin{cases} \mu_{\bar{x}} = \mu = 85 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{5}} \end{cases}$

May 2-8:59 AM

SG 20

Salaries of nurses are normally dist. with $\mu = \$7200/\text{mo.}$ and $\sigma = 300/\text{mo.}$

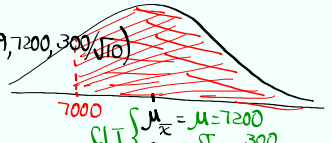
If we randomly select $n = 10$ 10 nurses, find $\bar{x}$ the Prob. that their mean monthly salary

a) is more than \$7000.

$P(\bar{x} > 7000)$

=normalcdf(7000, E99, 7200, 300/√10)

= .982



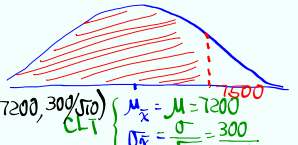

CLT $\begin{cases} \mu_{\bar{x}} = \mu = 7200 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{300}{\sqrt{10}} \end{cases}$

b) is less than \$7500.

$P(\bar{x} < 7500)$

=normalcdf(−E99, 7500, 7200, 300/√10)

≈ .999




CLT $\begin{cases} \mu_{\bar{x}} = \mu = 7200 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{300}{\sqrt{10}} \end{cases}$

May 2-9:18 AM

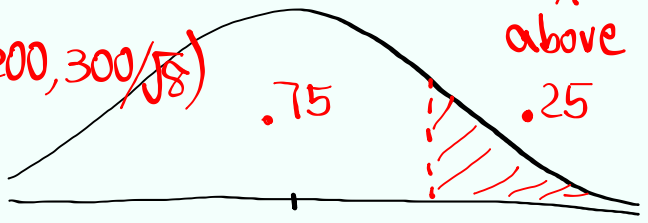
c) Find $\bar{x} = Q_3$ for randomly selected 8 nurses.

75% below, 25% above

$\bar{x} = \text{invNorm}(.75, 7200, 300/\sqrt{8})$

$= 7271.540$

$\approx \boxed{7272}$




CLT $\begin{cases} \mu_{\bar{x}} = \mu = 7200 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{300}{\sqrt{8}} \end{cases}$

May 2-9:27 AM

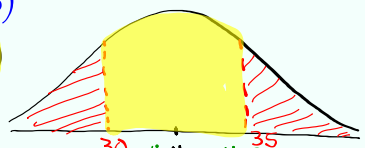
Ages of College students are normally dist. with $\mu = 32.5$ yrs and $\sigma = 6.5$ yrs.

If we randomly select $n = 2$ students, find the prob. that their mean age ($\bar{x}$) is below 30 yrs or above 35 yrs.

$P(\bar{x} < 30 \text{ OR } \bar{x} > 35)$

$= 1 - P(30 < \bar{x} < 35)$

↑ Total Area = Total Prob.




CLT $\begin{cases} \mu_{\bar{x}} = \mu = 32.5 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6.5}{\sqrt{2}} \end{cases}$

$= 1 - \text{normalcdf}(30, 35, 32.5, 6.5/\sqrt{2})$

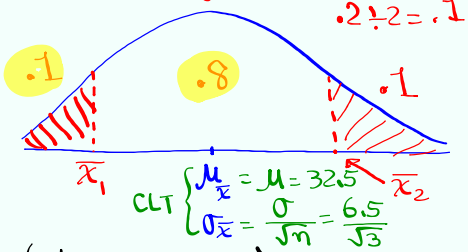
$\approx \boxed{.586}$

May 2-9:31 AM

Find two means for randomly selected group of 3 students that separate the middle 80% from the rest. Round to 1-dec. place.

$1-.8=.2$

$.2\div 2=.1$




CLT $\begin{cases} \mu_{\bar{x}} = \mu = 32.5 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6.5}{\sqrt{3}} \end{cases}$

$\bar{x}_1 = P_{10} = \text{invNorm}(.1, 32.5, 6.5/\sqrt{3})$

$\approx \boxed{27.7}$

$\bar{x}_2 = P_{90} = \text{invNorm}(.9, 32.5, 6.5/\sqrt{3})$

$= \boxed{37.3}$

SG 20

May 2-9:39 AM

$\alpha$ Alpha

$\alpha$ Significance level

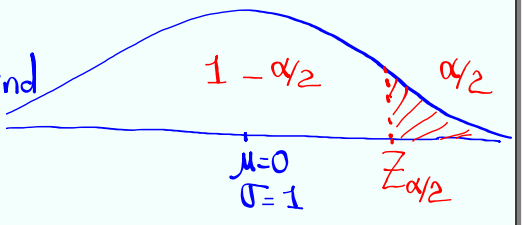
SG 21

$0 < \alpha < 1$

$\alpha/2$ is the area of the right-Tail of graph of Prob. dist.

$Z_{\alpha/2}$ is the critical value that separates the right-Tail area $\alpha/2$ from the rest.

we use invNorm to find $Z_{\alpha/2}$

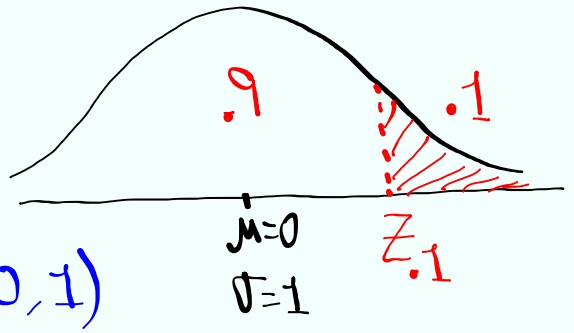



May 2-9:47 AM

Find $z_{.1}$

$\frac{\alpha}{2} = .1$ Right-Tail Area

$\alpha = .2$




$z_{.1} = \text{invNorm}(.9, 0, 1)$

$\approx \boxed{1.282}$

May 2-9:54 AM

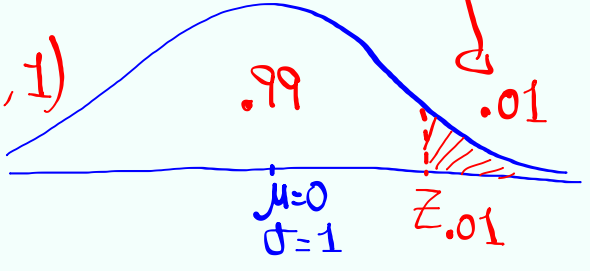
Find $z_{\alpha/2}$ for $\alpha = .02$.

$\alpha = .02$

$\alpha/2 = .01 \rightarrow$ Find $z_{.01}$

$z_{.01} = \text{invNorm}(.99, 0, 1)$

$\approx \boxed{2.326}$




May 2-9:57 AM

$\alpha$ Significance level

$0 < \alpha < 1$

$\alpha/2$ area of right tail

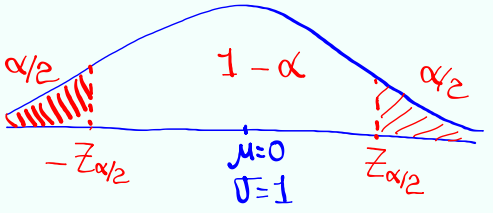
$1 - \alpha$ Middle area

$(1-\alpha)\cdot 100\%$ Confidence level

$\pm z_{\alpha/2}$ Critical values

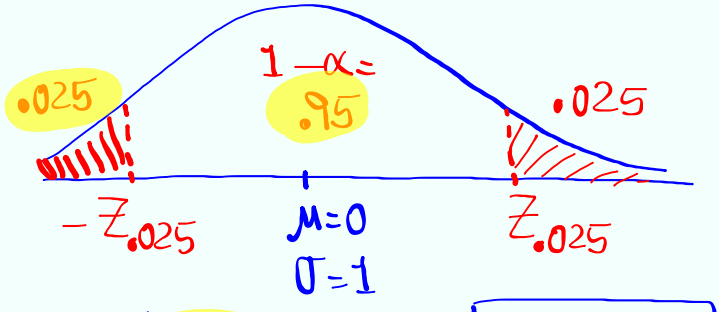
They separate the middle area of $1-\alpha$ from the rest.

We use invNorm to find $\pm z_{\alpha/2}$.




May 2-9:59 AM

Find $\pm z_{\alpha/2}$ for $\alpha = .05$.




$z_{.025} = \text{invNorm}(.975, 0, 1) \approx 1.960$

May 2-10:03 AM

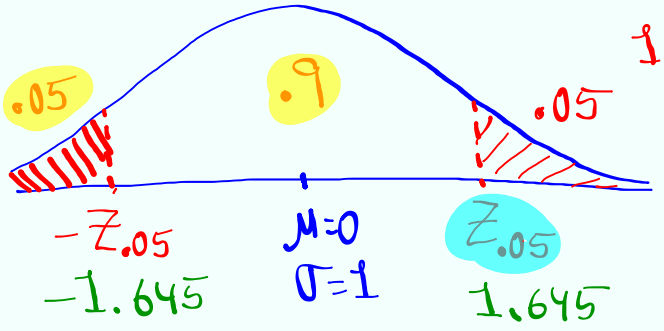
Find $\pm z_{\alpha/2}$ for 90% Conf. level.

Middle Area = .9

$1-\alpha = .9$

$\alpha = .1$

$\alpha/2 = .05$




$z_{.05} = \text{invNorm}(.95, 0, 1) = \boxed{1.645}$

May 2-10:07 AM

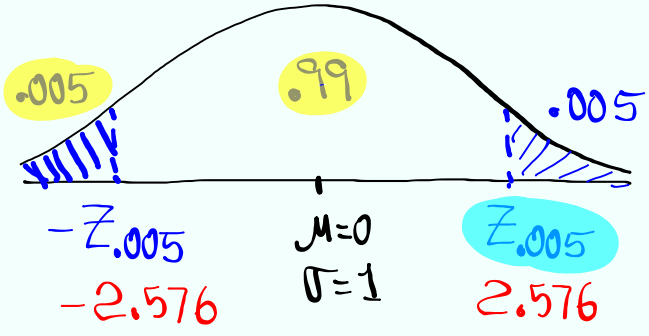
Find $\pm z_{\alpha/2}$ for 99% Conf. level. } Middle Area = .99

C-level }

$1-\alpha = .99$

$\alpha = .01$

$\alpha/2 = .005$




$z_{.005} = \text{invNorm}(.995, 0, 1) = \boxed{2.576}$

May 2-10:11 AM

Estimating Parameters

Parameters describe Population

Statistic describe Sample

we use Statistic as starting Point to estimate Parameter.

we use $\bar{x}$ to estimate $\mu$.

we use $s$ to estimate $\sigma$.

we use $\hat{p}$ to estimate $p$

$\hat{p}$ — p-hat — Sample Proportion

$p$ — Population Proportion

| To estimate | we use |
|---|---|
| $p$ | $\hat{p}$ |
| $\mu$ | $\bar{x}$ |
| $\sigma$ | $s$ |

Point-estimate

May 2-10:26 AM

when estimating parameters, the answer will be range of values

Confidence Interval

Probability that the parameter falls within the Confidence Interval is called Confidence level.

Middle Area

May 2-10:31 AM

Suppose the Conf. interval for Population mean is 70 to 90 with 98% Conf. level.

$$P(70 < \mu < 90) = .98$$

Suppose the Conf. interval for Pop. Prop. is between .34 to .48 with 90% Conf. level.

$$P(.34 < p < .48) = .9$$

May 2-10:34 AM

Estimating Population Proportion

$$\hat{P} - E < P < \hat{P} + E$$

$\hat{P}$ – P-hat, Sample Proportion

$P$ – Population Proportion

$E$ – Margin of error

$\hat{P} = \frac{x}{n}$ ← $x$: # of favorable responses; $n$: Sample Size

$\hat{q} = 1 - \hat{P}$

$$E = Z_{\alpha/2} \cdot \sqrt{\frac{\hat{P} \cdot \hat{q}}{n}}$$

$Z_{\alpha/2}$ – Critical value for $(1-\alpha)\cdot 100\%$ C-level.

May 2-10:38 AM

I Surveyed 100 students, 80 of them liked online classes.

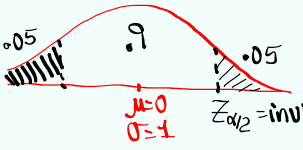
$n = 100$ $\quad \hat{p} = \frac{x}{n} = \frac{80}{100} = .8$

$x = 80$ $\quad \hat{q} = 1 - \hat{p} = 1 - .8 = .2$

Find 90% Conf. interval for the prop. of all Students that like online classes.

C-level: .9

$E = z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.645 \cdot \sqrt{\frac{(.8)(.2)}{100}} \approx .07$




$\hat{p} - E < p < \hat{p} + E$

$.8 - .07 < p < .8 + .07$

$.73 < p < .87$

I am 90% confident that between 73% and 87% of all students like online classes.

May 2-10:43 AM

Using TI:

STAT → TESTS ↓ 1-PropZInt

$.734 < p < .866$

$.73 < p < .87$

x: 80

n = 100

C-level: .9

Calculate

May 2-10:51 AM

I surveyed 80 students, 12 were smokers.

$n = 80$ $x = 12$

Find 99% Conf. interval for the prop. of all students that are smokers.

↳ C-level: .99

1-Prop ZInt

$x = 12$

$n = 80$

C-level: .99

$\hat{P} = \frac{.25 + .05}{2} = .15$

$.05 < P < .25$

I am 99% confident that between 5% & 25% of all students smoke.

$E = \frac{.25 - .05}{2} = .1$

May 2-10:55 AM

I surveyed 250 voters and 60% of them were in favor of immigration reform.

$n = 250$ $\hat{P} = .6$

$\hat{P} = \frac{x}{n}$ $x = n\hat{P}$

$x = 250(.6) = 150$

if decimal → Round-up

Find Conf. interval for the prop. of all voters in support of immigration reform.

↳ NO C-level

use 95%

$.54 < P < .66$

$\hat{P} = \frac{.66 + .54}{2} = .6$

$E = \frac{.66 - .54}{2} = .06$

1-Prop ZInt

$x = 150$

$n = 250$

C-level: .95

Calculate

May 2-11:01 AM