Statistical Inference Theory
Lesson 29
Estimating the Mean μ of a Population

29.1- What is the error created when using a point estimate?

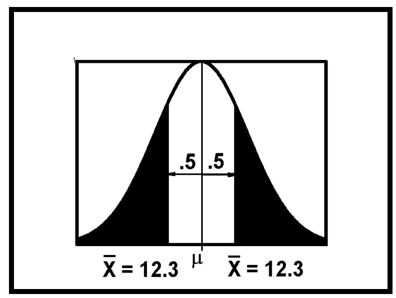
29.1 Problem 1:

➤(a).

It is almost certain that $\overline{X}$ is smaller or larger than μ. The difference between $\overline{X}$ and μ is the error

$e^* = \pm(\overline{X} - \mu)$.

a.



We need to find the probability that the error exceeds 0.5 ounces.

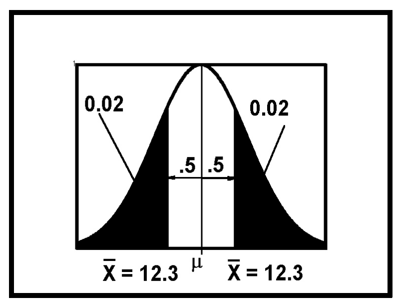
Step 1: Since we have the standard deviation of the population,

σ = 1.5 ounces.

Step 2: The sample size N = 50.

Step 3:

a.



$$e^* = \pm(\overline{X} - \mu) = \pm z\frac{\sigma}{\sqrt{N}} = \pm z\frac{1.5}{\sqrt{50}} \approx \pm z(0.212) = \pm 0.5 .$$

Step 4: Solving for z gives $z = \dfrac{0.5}{0.212} \approx 2.36.$

Step 5:

From the normal distribution table: $P\{ e^* > 0.5\} = 1 - 0.4909 - 0.4909 = 0.0182 \approx 0.02.$

➤(b).

Step 1: For the probability that the error will exceed 0.5 ounces is 0.01, we find z for the area 0.5 - 0.01/2 = 0.495: z = 2.58 .

Step 2: Since $e^* = \pm z\dfrac{1.5}{\sqrt{N}} = \pm(2.58)\dfrac{1.5}{\sqrt{N}} = \pm(2.58)\dfrac{1.5}{\sqrt{N}} = \pm\dfrac{3.87}{\sqrt{N}} = \pm 0.5$

Step 3: Solving for N gives $\sqrt{N} = 2(3.87) = 7.74$.

$N \approx 60$, minimum sample size.

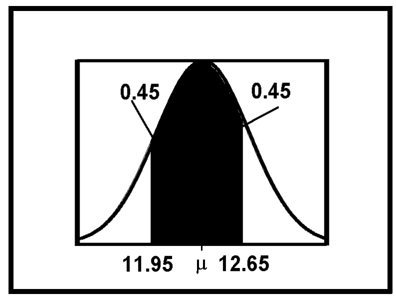
29.2 - What is the error created when using a Confidence Interval estimate ?

29.2 - Problem 1:

➤(a).

Since the confidence interval is 90%, we use the area $\frac{0.90}{2} = 0.45$ to find z = 1.64.

a.



Step 1:
$\overline{X} = 12.3$ and $\sigma = 1.5$

Step 2: N = 50

Step 3: $e^* = \pm z\frac{\sigma}{\sqrt{N}} = \pm 1.64\frac{1.5}{\sqrt{50}} \approx \pm 0.35$

Step 4: $\overline{X} - z\frac{\sigma}{\sqrt{N}} \le \mu \le \overline{X} + z\frac{\sigma}{\sqrt{N}}$

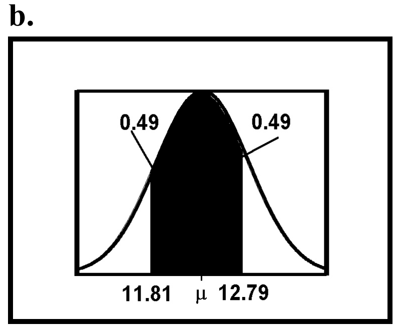
$12.3 - 0.35 \le \mu \le 12.3 + 0.35$ which gives $11.95 \le \mu \le 12.65$.

Step 5: The value for μ ranges between 11.95 and 12.65 with 90% probability.

➤(b).

Since the confidence interval is 98%, we use the area $\frac{0.98}{2} = 0.49$ to find z = 2.33.

b.



Step 1: $\overline{X} = 12.3$ and $\sigma = 1.5$

Step 2: N = 50

Step 3: $e^* = \pm z\frac{\sigma}{\sqrt{N}} = \pm 2.33\frac{1.5}{\sqrt{50}} \approx \pm 0.49$

Step 4: $\overline{X} - z\frac{\sigma}{\sqrt{N}} \le \mu \le \overline{X} + z\frac{\sigma}{\sqrt{N}}$:

$12.3 - 0.49 \le \mu \le 12.3 + 0.49$ which gives $11.81 \le \mu \le 12.79$.

Step 5: The value for μ ranges between 11.95 and 12.65 with 90% probability.

29.3 - Determining the Sample Size.

29.3 - Problem.1:

➤(a).

Since we want a confidence of 99%, we look up in the normal distribution table the area 0.495.

This gives $z = 2.57$.

From the statement of the problem, $e^* = 0.05$ and $\sigma = 0.3$.

Therefore, $N = \dfrac{z^2\sigma^2}{e^{*2}} = \dfrac{(2.57)^2(0.3)^2}{(0.05)^2} \approx 238$ students.

➤(b).

Since we want a confidence of 99%, we look up in the normal distribution table the area 0.495.

This gives $z = 2.57$.

From the statement of the problem, $e^* = 0.1$ and $\sigma = 0.3$.

Therefore,

$$N = \frac{z^2\sigma^2}{e^{*2}} = \frac{(2.57)^2(0.3)^2}{(0.1)^2} \approx 59 \text{ students.}$$

Supplementary Problems

1.

➤a.

Since we are using a fair coin, the chance of heads is $p = 0.50$. Therefore,

$\mu = Np = 100(.5) = 50$,

the true average number of heads that should appear. Since we need to compute the probability the error will be 10 heads or more, $e^* = \pm X - \mu = \pm 10$ heads or more.

This means that the deviation of X from μ is 10 heads or more from the expected number of heads $\mu = 50$.

Step 1: We use the formula $z = \frac{X - \mu}{\sigma}$.

Step 2: Since this a binomial distribution, $\sigma = \sqrt{Npq} = \sqrt{(100)(0.5)(0.5)} = 5$.

An error of 10 or more from the mean means that $X \geq 60$ or $X \leq 40$.

Step 3: For $X \geq 60$, we have $z = \frac{59.5 - 50}{5} = 1.90$.

Step 4: For $X \leq 40$, we have $z = \frac{40.5 - 50}{5} = -1.90$.

Step 5: Using the table for z = 1.90,

the area on both sides of the non-shaded is area is 0.4713 + 0.4713 = 0.9426.

Therefore, the probability that X will deviate from the true value μ is the shaded area
1 - 0.9426 = 0.0574 .

➤b.
Since X has a range of values from μ = 50, we need to find c^* where

$P\{\mu - e^* \leq X \leq \mu + e^*\} = P\{50 - e^* \leq X \leq 50 + e^*\} = 0.90$.

Step 1: Use the formula $e^* = z\sigma = z(5)$

Step 2:From symmetry, $P\{\mu \leq X \leq \mu + e^*\} = P\{\mu - e^* \leq X \leq \mu\} = \frac{0.90}{2} = 0.45$.

Step 3: Using the area portion of table C, z = 1.64.

Step 4: $e^* = z\sigma = 1.64(5) = 8.20$.

Step 5:Therefore, $P\{50 - 8.2 \leq X \leq 50 + 8.2\} = P\{41.80 \leq X \leq 58.20\}$

Step 6: Since the experiment results in number of heads, we should expect with 90% chance that the number of heads is between 42 and 58.

2.

➤a.
Since 60% of all medical students are female, p = 0.60. Therefore, μ = Np = 200(.6) = 120,

the average number of students that are female in the sample. Since we are interested in that X will deviate from $\mu = 120$ by 20 females or more, $e^* = \pm X - \mu = \pm 20$ females or more.

We use the formula $z = \dfrac{X - \mu}{\sigma}$.

Since this a binomial distribution, $\sigma = \sqrt{Npq} = \sqrt{(200)(0.6)(0.4)} \approx 6.93$.

An error of 20 or more from the mean means that $X \geq 140$ or $X \leq 100$.

For $X \geq 140$, we have $z = \dfrac{139.5 - 120}{6.93} \approx 2.81$.

For $X \leq 100$, we have $z = \dfrac{100.5 - 120}{6.93} \approx -2.81$.

Using table C for z = 2.81, the area on both sides of the non-shaded area is 0.4975 + 0.4975 = 0.995.

Therefore, the probability that x will deviate from the true value μ is the shaded area 1 - 0.995 = 0.005.

➤b.

Since x has a range of values from μ = 120, we need to find e^* where $P\{\mu - e^* \leq X \leq \mu + e^*\} = P\{120 - e^* \leq X \leq 120 + e^*\} = 0.95$.

Step 1: Use the formula $c^* = RVE = z\sigma = z(6.93)$ **.**

Step 2: From symmetry, $P\{\mu \leq X \leq \mu + e^*\} = P\{\mu - e^* \leq X \leq \mu\} = \dfrac{0.95}{2} = 0.475$.

Step 3: Using the area portion of the table, z = 1.96.

Step 4: $e^* = z\sigma = 1.96(6.93) = 13.58$.

Step 5: Therefore,

$P\{120 - 13.58 \leq X \leq 120 + 13.58\} = P\{106.42 \leq X \leq 133.58\}$

Step 6: With 95% chance, the number of students in the sample that are female should range between 107 and 133.

3.

➤a.

Step 1: Find the probability that a randomly selected bottle contains less than 16 ounces. Since the amount filled of each bottle is normally distributed, we have

$$z = \frac{16 - 16.2}{0.15} \approx -1.33.$$

From the normal distribution table,

$P\{z \le -1.33\} = 0.5 - 0.4082 \approx 0.09$.

Therefore, the probability that a bottle selected has less than 16 ounces is approximately 0.09.

Step 2: X is a binomial random distribution, where p = 0.09, N = 200.

We use the normal distribution to approximate the binomial distribution.

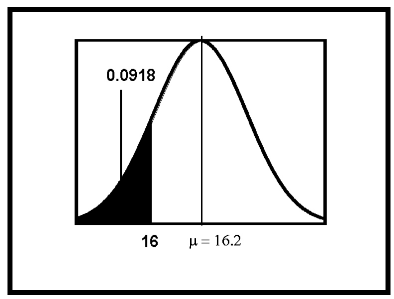
$\mu = pN = (0.09)(200) = 18.$

$$\sigma = \sqrt{Np(1 - p)} = \sqrt{(200)(0.09)(0.91)} \approx 4.05$$

$X = \mu \pm 5 = 18 \pm 5$

$$z = \frac{X - \mu}{\sigma} = \pm\frac{4.5}{4.05} \approx \pm 1.11$$

a.



From the normal distribution table,

$P\{z \ge 1.11\} + P\{z \le -1.11\} = (0.5 - 0.3665) + (0.5 - 0.3665) = 0.1335 + 0.1335 = 0.267$.

➤b.

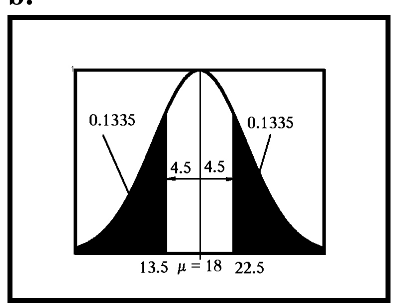
$X = \mu + z\sigma = 18 + z(4.05)$

For the area $\frac{0.90}{2} = 0.45$, $z = \pm 1.64$.

$X = \mu + z\sigma = 18 + z(4.05) = 18 \pm 1.64(4.05)$

Therefore, $P\{11 \le X \le 24\} \approx 0.90$

b.



4.

➤a.

Step 1: We need to find the probability that a random hand contains all black cards.
The number of possible hands:

$\binom{52}{5} = 2{,}598{,}960.$

The number of possible hands that contain only black cards: $\binom{26}{5} = 65{,}780.$

$$p = \frac{\binom{26}{5}}{\binom{52}{5}} \approx 0.025.$$

$\mu = Np = 1000(0.025) = 25$ hands.

➤b.

$$\sigma = \sqrt{Np(1-p)} = \sqrt{(1000)(0.025)(0.975)} \approx 4.94$$

$X = \mu \pm 5 = 25 \pm 5$

$$z = \frac{X-\mu}{\sigma} = \pm\frac{4.5}{4.94} \approx \pm 0.91$$

From the normal distribution table, $P\{z \geq 0.91\} + P\{z \leq -0.91\} = (0.5 - 0.3186) + (0.5 - 0.3186)$ $= 0.1814 + 0.1814 = 0.3628$.

➤c.

$X = \mu + z\sigma = 25 + z(4.94)$

For the area $\frac{0.95}{2} = 0.475$, $z = \pm 1.96$.

$X = \mu + z\sigma = 25 + z(4.94) = 25 \pm 1.96(4.94)$

Therefore, $P\{16 \leq X \leq 34\} \approx 0.95$.

5.

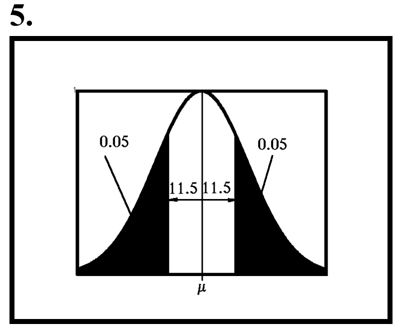
$X = \mu + z\sigma$

$$X = \mu + z\sqrt{Npq} = \mu \pm 11.5$$

$$z\sqrt{Npq} = \pm 11.5$$

$$\frac{0.90}{2} = 0.45$$

5.



$$\sqrt{Npq} = \pm\frac{11.5}{z}$$

From the normal distribution table, z = 1.64.

$$\sqrt{Npq} = \pm\frac{11.5}{1.64} = \pm 7.01$$

$$Npq = (7.01)^2 \approx 49.14$$

Since N = 1000 and q = 1 - p, it follows:

$$1000p(1 - p) = 49.14$$

$$1000[p - p^2] - 49.14 = 0$$

$$1000p^2 - 1000p + 49.14 = 0$$

Using the quadratic formula and since p and q are interchangeable,

$p \approx 0.95$.

$\mu = Np = 1000(0.95) = 950$

$$\sigma = \sqrt{Npq} = \sqrt{1000(0.95)(0.05)} = 6.89$$

6.
Step 1: $\mu - z\sigma_{\overline{X}} \le \overline{X} \le \mu + z\sigma_{\overline{X}}$

Step 2: Subtract μ for each side of the above inequality: $-z\sigma_{\overline{X}} \le \overline{X} - \mu \le z\sigma_{\overline{X}}$.

Step 3: Subtract $\overline{X}$ from each side of the above inequality: $-\overline{X} - z\sigma_{\overline{X}} \le -\mu \le \overline{X} + z\sigma_{\overline{X}}$.

Step 4: Multiply the above inequality by -1: $\overline{X} + z\sigma_{\overline{X}} \ge \mu \ge \overline{X} - z\sigma_{\overline{X}}$

7.
➤a.
To compute the average $\overline{X}$, we need to complete the following table:

Interest Rates on Home Mortgages	Number of Banks	col 1 x col 2
7.5%	24	
8.1%	55	
9.0%	11	
9.4%	10	
Total	100	

Multiplying column 1 and column 2 we get:

Interest Rates on Home Mortgages	Number of Banks	col 1 x col 2
7.5%	24	180
8.1%	55	445.5
9.0%	11	99
9.4%	10	94
Total	100	818.5

From the last column

$$\overline{X} = \frac{818.5}{100} \approx 8.18\%.$$

➤b.
Since this is a problem of descriptive statistics, we compute the standard deviation according to the method described in Lesson 3, Descriptive Statistics. To compute the standard deviation s, we need complete the following table:

Interest Rates on Home Mortgages	Number of Banks	(col 1 - $\overline{X}$)^2	col 2 x col 3
7.5%	24		
8.1%	55		
9.0%	11		
9.4%	10		
Total	100		

Since $\overline{X}$ = 8.18%, we have

Interest Rates on Home Mortgages	Number of Banks	(col 1 - $\overline{X}$)^2	col 2 x col 3
7.5%	24	0.47	11.28
8.1%	55	0.006	0.33
9.0%	11	0.67	7.37
9.4%	10	1.49	14.9
Total	100		33.88

Since N = 100, $s^2 = \frac{33.88}{100} \approx 0.34\%$

$s \approx \sqrt{0.34} \approx 0.58\%$.

Therefore, the stand error of the mean $\frac{s}{\sqrt{N}} = \frac{0.58\%}{10} = 0.058\%$.

➤c.
To find the probability that the error e* will exceed 0.1%, we use the error formula:

Step 1:$e^* = \pm\ z\frac{s}{\sqrt{N}} = \pm 0.1\%$.

Step 2:Since $\frac{s}{\sqrt{N}} = 0.058\%$,

we need to solve for z in the above formula

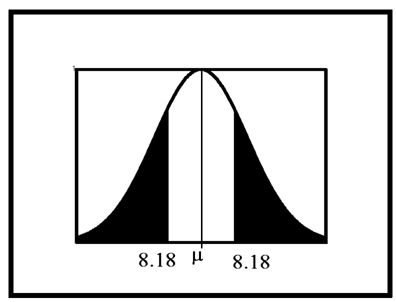
Step 3: z(0.058%) = ±0.1%.

Step 4: $z = \frac{0.1}{0.058} \approx \pm 1.72$. **fig. b**

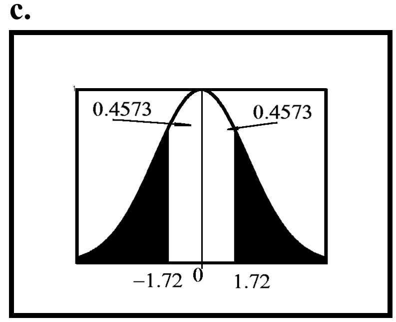
From the normal distribution table, the chance is

1 - 0.4573 - 0.4573 = 0.0854. **fig. c**

b.



c.



8.
➤a.
To compute the average $\overline{X}$, we need to complete the following table:

Number of pounds lost (rounded to nearest pound)	Number of members	COl 1 X COl 2
23	6	
35	10	
40	15	
45	5	
Total	36	

Multiplying column 1 and column 2 we get:

Number of pounds lost (rounded to nearest pound)	Number of members	COl 1 X COl 2
23	6	138
35	10	350
40	15	600
45	5	225
Total	36	1313

From the last column $\overline{X} = \frac{1313}{36} \approx 36.47$ pounds.

➤b.
To compute the standard deviation s, we need complete the following table:

Number of pounds lost (rounded to nearest pound)	Number of Members	$(\text{col } 1 - \overline{X})^2$	col 2 x col 3
23	6		
35	10		
40	15		
45	5		
Total	36		

Since $\overline{X} = 36.47$, we have

Number of pounds lost (rounded to nearest pound)	Number of Members	(col 1 - $\overline{X}$)^2	col 2 x col 3
23	6	181.44	1088.64
35	10	2.16	21.6
40	15	12.46	186.9
45	5	72.76	363.8
Total	36		1660.94

Since N = 36, $s^2 = \frac{1660.94}{36} \approx 46.14$ and $s \approx \sqrt{46.14} \approx 6.79$.

Therefore, the standard error of the mean $\sigma_{\overline{x}} = \frac{s}{\sqrt{N}} = \frac{6.79}{6} \approx 1.13$ pounds.

➤c.
To find the probability that the average weight of these 36 members is off by more than 1 pound from the true average of members that lost more than 10 pounds, we need the formula:

Step 1:$e^* = z\frac{s}{\sqrt{N}} = 1$

Step 2:Since $\sigma_{\overline{X}} = \frac{s}{\sqrt{N}} = 1.13$,

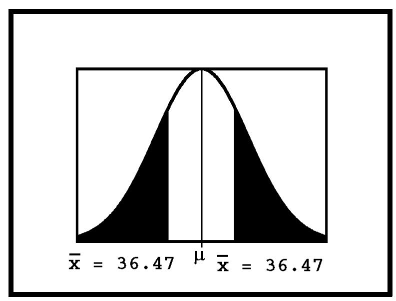
we need to solve for z in the above formula z(1.13) = 1.

Step 3: Therefore, $z = \frac{1}{1.13} \approx 0.88$.

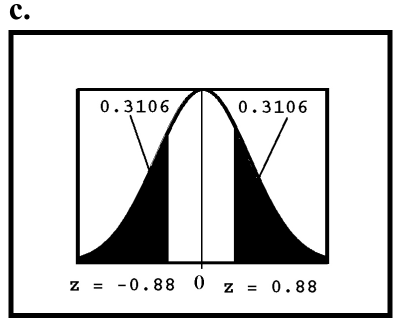
From the normal distribution table, the chance is

1- 0.3106 - 0.3106 = 0.3788.

c.



c.



9.

➤a.

To compute the average $\overline{X}$, we need to complete the following table:

Time it took to read novel (rounded to nearest minute)	**Number of readers**	**COl 1 X COl 2**
128	16	
150	34	
167	20	
205	15	
250	15	
Total		

Multiplying column 1 and column 2 we get:

Time it took to read novel (rounded to nearest minute)	**Number of readers**	**COl 1 X COl 2**
128	16	2048
150	34	5100
167	20	3340
205	15	3075
250	15	3750
Total	100	17313

From the last column $\overline{X} = \frac{17313}{100} = 173.13$

➤b.

To compute the standard deviation s, we need complete the following table:

Time it took to read novel (rounded to nearest minute)	**Number of readers**	**(col 1 -$\overline{X}$)^2**	**col 2 x col 3**
128	16		
150	34		
167	20		
205	15		

250	15		
Total			

Since $\overline{X}$ = 173.13, we have:

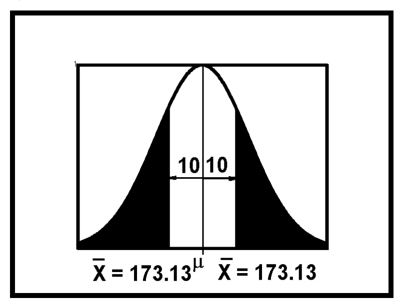
Time it took to read novel (rounded to nearest minute)	**Number of readers**	**(col 1 -$\overline{X}$)^2**	**col 2 x col 3**
128	16	2036.72	32587.52
150	34	535.00	18190.00
167	20	37.58	751.60
205	15	1015.70	15235.50
250	15	5909.00	88635.00
Total	100		155399.62

Since N = 100, $s^2 = \dfrac{155399.62}{100} \approx 1554$ and $s \approx \sqrt{1554} \approx 39.42$.

Therefore, the standard error of the mean is

$$\sigma_{\overline{X}} = \frac{s}{\sqrt{N}} = \frac{39.42}{10} \approx 3.94 \text{ minutes.}$$

b.



➤c.

Step 1: $e^* = z\dfrac{s}{\sqrt{N}} = 10$

Step 2: Since $\sigma_{\overline{X}} = \dfrac{s}{\sqrt{N}} = 3.94$,

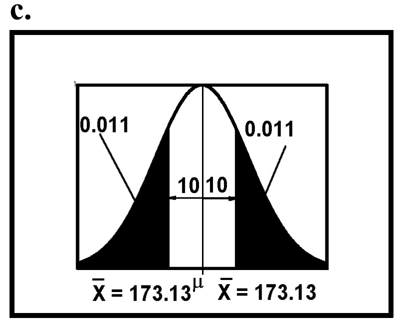
we need to solve for z in the above formula z(3.94) = 10.

Step 3: Therefore, $z = \dfrac{10}{3.94} \approx 2.54$.

From the normal distribution table, the chance is

1- 0.4945 - 0.4945 = 0.011 .

c.



10.

The determination of a minimum sample size is given by the formula

$N = \dfrac{z^2\sigma^2}{e^{*2}}$ where e^* is the desired error.

➤ **a**.

$$N = \frac{z^2\sigma^2}{e^{*2}}$$

$$N_1 = \frac{z^2\sigma^2}{(ae)^{*2}} = \frac{z^2\sigma^2}{a^2e^{*2}} = \frac{N}{a^2}$$

➤ **b**.

$$N_1 = \frac{N}{a^2} = \frac{100}{0.50^2} = 400$$

➤**c.**

$$N_1 = \frac{N}{a^2} = \frac{100}{1.10^2} \approx 83$$

11.

➤ **a**.

$$e^*_1 = \frac{z\sigma}{\sqrt{aN}} = \frac{z\sigma}{\sqrt{a}\sqrt{N}} = \frac{e*}{\sqrt{a}}$$

➤ **b**.

$$e^*_1 = \frac{0.1}{\sqrt{\frac{1}{2}}} = \frac{\sqrt{2}}{10} \approx 0.14$$

➤ **c.**

$$e^*_1 = \frac{0.01}{\sqrt{2}} \approx 0.007$$