Three main topics in statistics:

- Distribution
- Distribution of means
- Test statistic

Review #1: The central limit theorem describes three characteristics of the distribution of means. What does it say about its...

<table>
<thead>
<tr>
<th>shape</th>
<th>center</th>
<th>width</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Lectlet 8A Page 2

Review #1: The central limit theorem describes three characteristics of the distribution of means. What does it say about its...

Answer:

Yours

Mine

Review #2: Assume that the life of lightbulbs is normally distributed with mean $\mu = 1000$ hours and standard deviation $\sigma = 100$ hours. Sketch the distribution of lightbulb life on a piece of scratch paper. Then answer the following question: The "middle 95%" of lightbulb life ranges from about ______ hours to ______ hours.
Review #2: Assume that the life of lightbulbs is normally distributed with mean $\mu = 1000$ hours and standard deviation $\sigma = 100$ hours. Sketch the distribution of lightbulb life on a piece of scratch paper. Then answer the following question: The "middle 95%" of lightbulb life ranges from about ______ hours to ______ hours.

Answer:

![Distribution Diagram]

NB: Most of the time the value of the variable lies within about two standard deviations of the population mean.

Review #3: Assume that these lightbulbs are packed 100 to a carton, and that the 100 bulbs in a carton can be considered random samples from the lightbulb population. Sketch the distribution of carton mean bulb life on a piece of scratch paper. Then answer the following question: The "middle 95%" of carton mean life ranges from about ______ hours to ______ hours.

Yours | Yours
--- | ---
Mine | Mine
Review #3: Assume that these lightbulbs are packed 100 to a carton, and that the 100 bulbs in a carton can be considered random samples from the lightbulb population. Sketch the distribution of carton mean bulb life on a piece of scratch paper. Then answer the following question: The "middle 95%" of carton mean life ranges from about ______ hours to ______ hours. 
Answer:

\[
\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{100}} = \frac{100}{10} = 10
\]

Mean life (hours) (Samples of size 100)

NB: Most of the time the sample mean lies within about two standard errors of the population mean.

New material begins...

"Most" = "95%", but what does "about two" mean?
Proportions of areas under the normal curve (excerpt)

<table>
<thead>
<tr>
<th>Area between mean and z</th>
<th>Area beyond z</th>
</tr>
</thead>
<tbody>
<tr>
<td>z</td>
<td>(A)</td>
</tr>
<tr>
<td>1.95</td>
<td>.4744</td>
</tr>
<tr>
<td>1.96</td>
<td>.4750</td>
</tr>
<tr>
<td>1.97</td>
<td>.4756</td>
</tr>
<tr>
<td>1.98</td>
<td>.4757</td>
</tr>
<tr>
<td>1.99</td>
<td>.4758</td>
</tr>
</tbody>
</table>

Therefore "about two" = "1.96," so

95% of the time the sample mean lies within 1.96 standard errors of the population mean.
95% of the time the sample mean lies within 1.96 standard errors of the population mean.

We reverse that statement to form the "confidence interval":

95% of the time the population mean lies within 1.96 standard errors of the sample mean, even if you don't know the magnitude of the population mean.

Let's call "1.96" the "critical value of \( z \)". Then, we can transform the above statement into

<table>
<thead>
<tr>
<th>95% of the time</th>
<th>the unknown magnitude of the population mean is less than 1.96 standard errors above the sample mean</th>
<th>and greater than 1.96 standard errors below the sample mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>or...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% of the time</td>
<td>the unknown magnitude of ( \mu )</td>
<td>( \mu &lt; \bar{X} + z_{cv} \sigma_{\bar{X}} ) \quad and ( \mu &gt; \bar{X} - z_{cv} \sigma_{\bar{X}} )</td>
</tr>
</tbody>
</table>

(Recall the terminology:

- the population mean is \( \mu \)
- the sample mean is \( \bar{X} \)
- the standard error of the mean is \( \sigma_{\bar{X}} \))

Combining the two inequalities for \( \mu \) gives the standard form of the confidence interval:

95% of the time, \( \bar{X} - z_{cv} \sigma_{\bar{X}} < \mu < \bar{X} + z_{cv} \sigma_{\bar{X}} \)
The confidence interval:
Where's the (unknown) population mean?
Probably within about two standard errors of the sample mean.

\[ \bar{X} - z_{cv} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{cv} \frac{\sigma}{\sqrt{n}} \]

To recapitulate...
We don't know \( \mu \).
\( \bar{X} \) is our best point-estimate of \( \mu \), but it's very unlikely that \( \mu \) is exactly equal to \( \bar{X} \).

So we agree to define an interval which very likely contains \( \mu \). We call that interval the confidence interval.

*End of lectlet.*