Section 6.1 - Confidence Intervals for the Mean (Large Samples)

1. Definitions:

   (a) Inferential statistics: using sample statistics to estimate the value of an unknown parameter.

   (b) **Point estimate**: a single value estimate for a population parameter. (i.e. $\bar{x}$ estimates μ)

   (c) **Interval estimate**: an interval (range of values) used to estimate a population parameter. Although the point estimate is not likely exactly μ, it's probably close (margin of error).

   (d) **Level of confidence** c: the probability that the interval estimate contains the population parameter.

   (e) **Critical values**: (Assume $n \geq 30$, so CLT applies) c is the area under the standard normal curve between the **critical values**, $-z_c$ and z_c. (Look at for 90%)

   (f) The difference between the point estimate and the actual parameter value is called the **sampling error** $(= \bar{x} - \mu)$.

2. The reasoning of statistical estimation:

   (a) To estimate an unknown population mean μ, use the mean $\bar{x}$ of a random sample.

   (b) We know that $\bar{x}$ is an unbiased estimate of μ. The law of large numbers says that $\bar{x}$ will always get close to μ if we take a large enough sample. Nonetheless, $\bar{x}$ will rarely be exactly equal to μ, so our estimate has some error.

   (c) We know the sampling distribution of $\bar{x}$. In repeated samples, $\bar{x}$ has the Normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$.

   (d) The 95 part of the 68-95-99.7 rule for Normal distributions says that $\bar{x}$ and its mean μ are within two standard deviations of each other in 95% of all samples.

   BIG IDEA: The sampling distribution of $\bar{x}$ tells us how big the error is likely to be when we use $\bar{x}$ to estimate μ. When we know the population standard deviation σ, this distribution is Normal, has known standard deviation, and is centered at the unknown population mean μ.

3. Common Confidence Intervals: (Note: 95% c.l. is slightly different than 2 from 68-95-99.7)

Confidence level C	90%	95%	99%
Critical value z^*	1.645	1.960	2.576

4. Margin of error:
   Given a confidence level c, the m.o.e. E is the greatest possible distance between the point estimate and the value of the parameter it is estimating.

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

   When $n \geq 30$, s can be used in place of σ.

5. Confidence Interval for the Mean of a Normal Population: ($-z_c$ is inverse Normal $\frac{1-c}{2}$)
   A c-confidence interval for μ is

$$\bar{x} - E < \mu < \bar{x} + E.$$

   The probability that the confidence interval contains μ is c.

6. Example #1:
Here are the IQ test scores of 31 seventh-grade girls in a Midwest school district:

114	100	104	89	102	91	114	114	103	105	
108	130	120	132	111	128	118	119	86	72	
111	103	74	112	107	103	98	96	112	112	93

Treat the 31 girls as an SRS of all seventh-grade girls in the school district. Suppose that the standard deviation of IQ scores in this population is known to be $\sigma = 15$. Give a 99% confidence interval for the mean score in the population. (Note: $\bar{x} = \frac{3281}{31} \approx 105.85$.)

7. All other things being equal, the margin of error $\left(z_c \frac{\sigma}{\sqrt{n}}\right)$ of a confidence interval gets smaller as

   - the confidence level c decreases (z_c decreases),
   - the population standard deviation σ decreases, and
   - the sample size n increases.

8. Sample Size for Desired Margin of Error:
The confidence interval for the mean of a Normal population will have a specified margin of error m when the sample size is

$$n = \left(\frac{z_c \sigma}{E}\right)^2.$$

9. Example #2:
How large a sample of schoolgirls from Example #1 would be needed to estimate the mean IQ score μ within ± 5 points with 99% confidence?