Estimated Margin of Error for $\bar{x}$. Consult pp. 62-65. This follow-on to lecture 1-12-09 extends the margin of error concept to the sample average $\bar{x}$. I am again considering the case in which $n$ and $N-n$ are both "large enough" so that a bell curve approximation is justified. I will not introduce Student's-t at this point, so the readings of pp. 586-606 are not particularly relevant.

A Question: What is the average number $\mu$ of pictures or graphics per page from pp. 1 to 767 of our textbook? This "population average" $\mu$ (Greek, pronounced "mu") is ordinarily estimated by the sample average $\bar{x}$. What is the estimated margin of error in this setup?

I've randomly sampled 36 pages by perusing the table of random digits A-94, skipping over those outside the range 001 to 767 (consult the table). Those in bold have graphics and I've indicated the number of graphics in parentheses (graphic sub-components of a graphic display are not counted individually. It is a serious point, needing clarification for serious work, but let's just suppose that we'recounting graphics drawing attention to markedly different parts of the page.

716, 032(1), 463 473, 200(1), 731 727(1), 039, 759 (skip 944) 043 (skip 890 and 877) 764(3), 364(1), 132(1) 512, 678, 098(3) 181(3), 027(1), 133 622(2) (skip 922) 666(1), 310(1) (skip 844) 720(1) (skip 945) 639, 112(5), 285(1) (skip duplicate 112) 429, 471(1) (skip duplicate 112) 647(1) (skip 770) 183(1), 071, 359 412, 585(1), 428 042(2)

Here are resulting scores $x =$ number of graphics on each page of 36 sample pages:

\{0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 3, 1, 1, 0, 0, 3, 3, 1, 0, 2, 1, 1, 0, 5, 1, 0, 1, 1, 0, 0, 0, 1, 0, 2\}
1a. The sample average number of graphics per page is

\[
\bar{x} = \frac{\sum x}{n} = \frac{1+1+1+3+1+1+3+3+1+2+1+1+1+5+1+1+1+1+1+2}{36}
\]

\[
= \frac{16(0)+14(1)+2(2)+3(3)+1(5)}{36}
\]

\[\sim 0.888889.\]

The sample standard deviation \(s\) (see page 64) is computed

\[
s = \sqrt{\frac{16(0-0.888889)^2 + 14(1-0.888889)^2 + 2(2-0.888889)^2 + 3(3-0.888889)^2 + 1(5-0.888889)^2}{36-1}}
\]

\[\sim 1.14087\]

1b. Point estimate of \(\mu = \bar{x} \sim 0.888889\) and sample standard deviation \(s \sim 1.14087\).

Estimated margin of error of the estimator \(\bar{x}\) is calculated as follows:

\[
1.96 \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = 1.96 \frac{1.14087}{\sqrt{36}} \sqrt{\frac{767-36}{767-1}}
\]

\[\sim 0.36407\]

1c. Claim made for estimated margin of error:

Around 95% of random samples of \(n = 36\) pages from the book of \(N = 767\) pages will produce an interval

\(\bar{x} \pm \) estimated margin of error

that will cover the actual value of \(\mu = \) average number of pictures/graphics per page of pages 1 through 767 of the textbook.

Our random sample of \(n = 36\) pages from the \(N = 767\) pages of the book has produced the interval

\(\bar{x} \pm \) estimated margin of error

\[= 0.888889 \pm 0.36407\]

\[= [0.5248196, 1.25296]\]

This interval is called a "95% confidence interval for \(\mu\)."
1d. Around 95% of samples of 36 are "good samples" whose 95% confidence interval $\bar{x} \pm \text{estimated margin of error}$ covers the true value of $\mu$. Is our sample of 36 a "good one?" We don't know. Nonetheless, the method of reporting the sampling results does give a pretty good idea of the reliability of the findings. For this 95% chance to apply we require that $n$ is "large enough" and so too is $N-n$. At this point we're not putting too fine a point on it. Consult the readings.

2. Other confidence levels. Consult Table T page A-97. Consult the next to bottom row of this table (it has an $\infty$ symbol at the left of that row, denoting large sample size). Notice the entry 1.960 of this row and that directly below this is confidence level 95%. Likewise, to obtain a 99% confidence level in the large sample confidence intervals for either $p$ or for $\mu$ you would substitute 2.576 ro 1.96 (see the final entries in the last two rows of the table). So the 99% confidence interval for $\mu$ takes the form

$$\bar{x} \pm (2.576 \text{ not } 1.96) \frac{s}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Claim: Around 99% of samples of $n$ from $N$ will produce an interval covering the population mean $\mu$. This requires random sampling and that both $n$, $N-n$ are "large enough."

Later on we'll learn the t-table. For now, it is a convenient place to find the correct $z$-value in order to have a confidence interval for confidence levels .8, .9, .95, .98, .99.