Math 343, Ian Chris - Phillip - "Course Assistant" Stewart McGregor.
Office hours: in web page M Tu 3-4 Th 11-12. *(Will probably change to M Tu 3-9 Th, 11-12)*
or by appointment. *Separate Zoom meeting* *(Will put this and course url page in Black)*
Prove: Math 252 or equivalent. *(reading web page 1-17, but tough; lack of recognition this)*
Suggested similar to Math 243, but better theoretical background. Will stay at little bit after class.

Notation: *wording (see some links on course page).*
- Will put videos, pdfs if notes. See class email messages.
- [Webmail reader]

End of note should be plain text.

For file submission (HW, exam, etc): pdf (e.g., Tex); his word files.

Single file per assignment.

Leave margins in scan (½ with). With resolution, near edge to near white, reasonable size.

File names: no spaces. Use only a-z, A-Z, 0-9, underscore.

Name on every page.

For this week: Read Ch 1 of book, Ch 2. *(will stay very little about Secs. 2.1-2.9)*

---

What is wrong with this pie chart?

Motivation: A simple random sample (SRS) of residents of Megalopolis (pp. 23, 682, 951) of size 986 is asked if they have had at least one coronavirus vaccination shot, 367 say yes, 619 no. Want to estimate what percentage of people had at least one shot.

Estimate is \( \frac{367}{986} \). (check 37.22%).
Will not be exactly right; each year's output will be different with a different sample.

Some terminology: the sample proportion (often called \( p \)) is a statistic (anything computed from a sample). The true proportion (often called \( \pi \)) is a parameter (any true description of the population). There will be an error \( |p - \pi| \) (depending on sample). How big is it likely to be?

Warning: statistical analysis makes assumptions which may not really be true.

Say the sample above was originally supposed to be 1100 people, but 114 of them couldn't be reached. Of those we were reached, 23 of them lied. Was our original sample of 1100 really a SRS? (For ex., did miss the Wang family, who moved there last week?)

What deal with that here.

Example. Measure salt concentration in a lake in g/liter. Different measurements give values 302, 312, 307. Suppose true concentration is 305. “Mean error” (MA, for mean absolute) is \( \frac{|-3| + 17 + 12|}{3} = 4 \) g/liter. Hard to work with theoretically. Instead, usually \( \sqrt{\frac{(-3)^2 + 7^2 + 2^2}{3}} = \sqrt{\frac{62}{3}} \approx 4.546 \text{ g/l} \)

General fact: \( MA \leq RMS \). Reason: for just two things. Want

\[ \frac{|x| + |y|}{2} \leq \sqrt{\frac{x^2 + y^2}{2}} \]

To see that this is true.

\[ \frac{x^2 + y^2}{2} - \left( \frac{|x| + |y|}{2} \right)^2 = \frac{x^2}{2} + \frac{y^2}{2} - \left( \frac{x^2}{4} + \frac{|x| |y|}{2} + \frac{y^2}{4} \right) \]

\[ \text{need this to be } \geq 0 \]

\[ \frac{1}{4} x^2 - \frac{|x| |y|}{2} + \frac{|y|^2}{4} = \frac{1}{4} (|x| - |y|)^2 \geq 0 \]

In Ch 1 (see for ex. page 7, or HW): \( 24.3 \pm 0.9 \) miles. Interpretation of 0.9 is “standard error” \( SE \), which is the least (notional) of RMS error as sample size \( \to \infty \) in a sufficiently random way (more later).

Std error didn't add; instead, say, if errors are independent:

\( 11.2 \pm 0.1 \) add \( 11.5 \pm 0.2 \) give \( 22.7 \pm \sqrt{(0.1)^2 + (0.2)^2} \) (not \( 11.61 \pm 0.3 \))