THIS STATISTICAL MODELS FOR MEAN

READ: FPP CH. 20, 21

MEASUREMENT ERROR, REVISITED: WHY DOES IT HELP TO TAKE AN AVERAGE?

IF NO BIAS

\[ (y_1)^{obs.1} = \text{TRUTH} + \text{ERROR}_1 \]
\[ (y_2)^{obs.2} = \text{TRUTH} + \text{ERROR}_2 \]
\[ \vdots \]
\[ (y_n)^{obs.n} = \text{TRUTH} + \text{ERROR}_n \]

MEAN
\[ \bar{y} = \text{TRUTH} + \text{AVERAGE OF RANDOM ERRORS} \]

AVERAGE OF N OBS

CANCELLATION OF +4 - TERMS

TYPICAL ERROR IN ANY SINGLE OBSERVATION IS ABOUT 0 IN SIZE, BUT TYPICAL SIZE OF AVERAGE
Of a random error will not be 0 but \( \text{SE}(\bar{y}) = \frac{\sigma}{\sqrt{n}} \) (a lot smaller than \( \sigma \) if \( n \) is large).

If bias

\[
\text{Obs}_1 = \text{Truth} + \text{Bias} + \text{Error}_1
\]

\[
\vdots
\]

\[
\text{Obs}_n = \text{Truth} + \text{Bias} + \text{Error}_n
\]

\[
\bar{y} = \text{Truth} + \text{Bias} + \text{Ave. of n random error}
\]

If bias, as \( n \) ↑ \( \bar{y} \) gets closer not to truth but to \( \text{Truth + Bias} \) & we can't make bias go away just by getting more data (this explains literary digest poll).
**Quantitative Summary**

**Quantity of Interest**
- Mean Amount of Money Owed on Pop. Waybills

**Estimate**
- \(ar{y} = 28.09\)

**Give or Take**
- \(SE(\bar{y}) = \frac{s}{\sqrt{n}} = 0.69\)
NORMAL CURVE: IMPOSSIBLE HERE BECAUSE IT PREDICTS A LOT OF VARIABLES WITH NEGATIVE $\mu$; LONG TAIL EVEN WORSE; SO HIST. HAS TO HAVE LONG RIGHT HAND TAIL.
\[ E_{\text{IID}}(\bar{x}) = \mu \] (Formula Sheet)

\[ \text{Estimate:} \]
\[ \text{SE}_{\text{IID}}(\bar{x}) = \frac{\sigma}{\sqrt{n}} \]

\[ \text{Nice but unusable (\sigma not known)} \]
\[ \text{SE}_{\text{IID}}(\bar{x}) = \frac{5}{\sqrt{10}} = 0.69 \]

\[ \text{SE \& 0.69} \]

\[ \text{LONG RUN histograms of } \bar{x} \]

\[ \text{(CLT)} \]

\[ \mu \]

\[ \text{EV} \]

Interpretation: I think \( \mu \) is around \( \$28.09 \) \( \bar{x} \), give or take about \( \$0.69 \) (SE of \( \bar{x} \)).