QMS 102 Measures of Variability

Coefficient of Variation

Are the size of Elephants “more variable” than the size of grasshoppers?
In absolute terms yes!

But we might want to make this decision with the effect of the mean removed?

Let's put the elephants on a diet (or fatten up the grasshoppers) until each group has the same mean size then see!
QMS 102  Coefficient of Variation
Now are the sizes of Elephants “more variable” than the sizes of grasshoppers when the means are “made” equal? The answer is not obvious!
QMS 102  Coefficient of Variation

Instead of fattening up the grasshoppers and/or putting the elephants on a diet we can achieve the same result by dividing every value in the populations by its mean.

In 1989 a typical house sold for $100,000 ± $10,000  \( R_{89}=20 \)

In 1996 a typical house sold for $200,000 ± $15,000  \( R_{96}=30 \)

For both the above, divide all values by the mean.

\( R^*_{89} = .2 \)

\( R^*_{96} = .15 \)

No units!
QMS 102  Coefficient of Variation

In the same way we can remove the “effect” of the mean on the standard deviation by dividing by the mean and expressing the standard deviation as a proportion of the mean.

Eg #1 Elephants   Data:  \( n = 10 \quad \bar{x}_e = 12,000 \text{ kg} \quad s_e = 2,000 \text{ kg} \)
Grasshopper Data:  \( n = 25 \quad \bar{x}_g = 51.0 \text{ g} \quad s_g = 21.0 \text{ g} \)

When a "mean free" measure of variability is required, ie to compare homogeneity, then use the Coefficient of Variation:

\[
CV = \frac{\sigma}{\mu} \quad \text{or} \quad CV = \frac{s}{\bar{x}}
\]

\[
CV_{\text{elephants}} = \frac{2000}{12000} = .17 < CV_{\text{grasshoppers}} = \frac{21}{51} = .41
\]

The coefficient of variation is often used to compare the variability of two or more groups!
Question:

Determine whether the waiting times with the “Old Policy” are more homogeneous (relatively less variable) than with the “New Policy”?

\[
CV \text{ “first 4 weeks”} = \frac{\sigma}{\mu} = \frac{6.93}{10.75} = 0.64 \text{ or } 64\%
\]

\[
CV \text{ “last week”} = \frac{6.58}{8.6} = 0.77 \text{ or } 77\%
\]

* we are treating this data as a population but in reality it is obviously a sample. Later we will be able to compute \( s_x \).