Raw Frequency vs Relative Frequency(%)  
\[ \frac{\text{how many}}{\text{how many} \div n} \times 100\% \]

L-16

Raw/Rel \rightarrow \text{histograms will be identical}

Density Scale: relative freq = area under histogram

Convention: all hist. sketches will be on density scale in this class

L-18 ex. 1 row for each nest (subject)
1.1 1 col for each (variable) location

\[
\begin{align*}
\text{Vine} \\
\text{Edge} \\
\text{Tree}
\end{align*}
\]

Variable: qual, nominal, not dich hist? no bar? yes

\[ n = 50 + 60 + 40 + 49 \]

1.2 1 row for each fish
1 col for each pigmentation class

Variable: qual, ordinal, not dich hist? no bar? yes

1.3 1 row for each litter
1 col for each size

Variable: quant, discrete, ratio hist? yes

\[ n = \text{sum of freq.} \]
1 row for each plant
1 col for each # of aphids

n = 424

Variable: quant, discrete, ratio

too much noise to get clear picture of shape

42 bars is too many # of aphids

too few bars, can't see shape

Histogram Shape

A and B different center, same spread + shape
A and C same center, same shape, different spread
A and D same center + spread, different shape

Symmetric

Non-Symmetric

left skew
right tail

right skewed
positive skewed

left skewed
negative skewed

Modes

bi-modal (multi-modal)
unimodal

Mode = point of highest frequency

If symmetric - mode = median = pt of symmetry

IF UNIMODAL
Median: $\sim y = 50^{th}$ percentile, 20th quantile

\[
\begin{align*}
Y_1 \left[ \begin{array}{c} 1 \\
2 \\
9 
\end{array} \right] & \quad \text{mean} \\
Y_2 \left[ \begin{array}{c} -3 \\
-2 \\
+5 
\end{array} \right] & \quad \text{outlier} \\
Y_n \left[ \begin{array}{c} a 
\end{array} \right] & \quad \text{mean}
\end{align*}
\]

Mean: balance point

Symmetric Unimodal

Move 1 end

Asymmetric Distribution: ask if mean or median
Spread measured by deviation

Deviations from mean: $y_n - \bar{y}$

\[
\begin{align*}
\left[ \frac{1}{n} \sum_{i=1}^{n} \right] & \quad \text{mean} \\
\left[ \frac{-3}{2} \right] & \quad \text{abs} \\
\left[ \frac{3}{5} \right] & \quad \text{mean} = \frac{16}{5}
\end{align*}
\]

\[
\left| y_n - \bar{y} \right|
\]

M.A.D. = $\frac{1}{n} \sum_{i=1}^{n} \left| y_i - \bar{y} \right|$

Mean Absolute Deviation

\[
\begin{align*}
\left[ \frac{1}{4} \right] & \quad \text{sub} \\
\bar{y} & \quad \text{square} \\
\left[ \frac{-3}{2} \right] & \quad \text{mean} = \frac{38}{3}
\end{align*}
\]

\[
\left[ \frac{26}{26} \right] \quad \text{in but... this squares units...}
\]

\[
\left[ \frac{1}{n-1} \right] \quad \text{sample variance} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2
\]

\[
\text{SD} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2} = s
\]

Sample Standard Deviation

(11)