3.1 Scatterplots & Correlation

Response variables measure the outcome of a study. An **explanatory variable** may help explain the changes in a response variable.

Ex: Determine the explanatory & response variable:
Adult volunteers drank a different number of cans of beer. 30 minutes later, their blood alcohol level was checked.

explanatory: # of cans of beer

response: blood alcohol level

Scatterplots show the relationship between 2 quantative variables measured on the same individual. Explanatory variable is typically graphed on the x-axis, responsevaridble is on the y-axis. Each individual is graphed as a point.

Examining a scatterplot (DOFS)

① Direction: if increasing from left to right, that is a positive association. if decreasing from left to right, that is a negative association.

② outliers: points (individuals) that don't fit the overall pattern

③ Form: linear or nonlinear (curved). Cluste

④ Strength: how close the data follows a straight lin pattern. weak, moderate, strong

Ex: SEC Foot ball example

Strong linear positive association, with no outliers

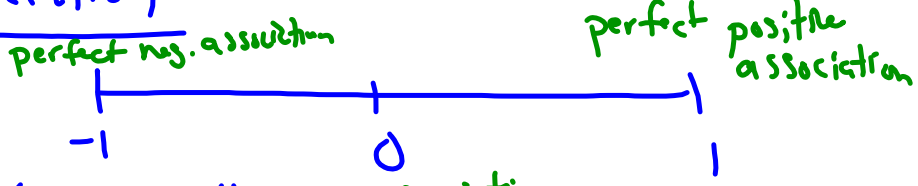
p. 149 check your understanding

moderate linear positive association no outliers
There are two clusters with a low duration interval and with a high duration interval

Correlation: r, measures the direction and strength of the linear relationship between 2 quantative variables.

Calculating correlation

① $-1 \leq r \leq 1$



② If r is positive = positive association
If r is negative = negative association

③ only measures linear relationships

④ unitless

Ex: p153 check your understanding

b) $r = .7$

strong linear positive association between storms predicted & storms observed.

d) $r = -.1$

weak linear negative association between last yrs % return & this years % return.

Properties of Correlation

1. Correlation makes no distinction between explanatory and response variable.
2. r doesn't change when we change units of measure.
3. Correlation does NOT imply causation.
4. Must have 2 quantitative variables.
5. Only measures strength of linear relationships, not curved. So..... GRAPH YOUR DATA FIRST!!
6. Values of r close to -1 and 1 do NOT garauntee a linear relationship between 2 variables (see p. 156 for a picture of this).
7. r is not a resistant measure so be careful when using it when outliers are present. Must graph data first to see this.
8. Correlation is not a complete summary of 2 variable data. You should also give means and standard deviations of both variables.