

2nd Week Summary (01/24/25)

- Relationship between two variables :

- Explanatory and response variables

Categorical vs. Quantitative : side-by-side box-plot

Quantitative vs. Quantitative : correlation $r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$

- Properties of the Correlation r:

Takes values between -1 and 1

$r = 1$ or $r = -1$ implies that the points lie on a straight line

$r = 0$ implies that there is no **linear** association

$r < 0$ implies that there is a negative **linear** association & $r > 0$ implies positive **linear** association

If the x and y variables are switched, the correlation will stay the same

r does not change when we change the units of measurement of x, y , or both.

r is strongly affected by a few outliers.

Chapter 3: Probability

- The **probability** of any event A of a random phenomenon is the proportion of times the event would occur in a very long series of repetitions.

- The **sample space** S of a random phenomenon is the set of all possible outcomes.

- An **event** is a subset of the sample space.

- Probability Rules

The probability $P(A)$ of any event A satisfies $0 < P(A) < 1$.

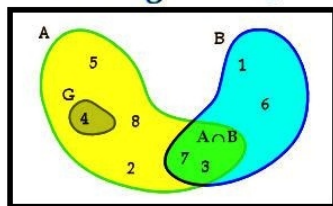
If S is the sample space, then $P(S) = 1$.

Two events A and B are **disjoint** if they have no outcomes in common and so can never occur together. If A and B are disjoint, $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B)$

For any event A , $P(A \text{ does not occur}) = P(\bar{A}) = 1 - P(A)$.

- Addition rule in general : $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Venn Diagram set operations



$$A = \{5, 8, 2, 4, 7, 3\}$$

$$B = \{1, 6, 7, 3\}$$

$$G = \{4\}$$

$$A \cap B = \{7, 3\}$$

$$A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

$$G \cap B = \emptyset$$

Tree Diagram

