1) Core Definitions

- **Mutually Exclusive (Disjoint)** — cannot occur simultaneously
  $P(A \cap B) = 0$
- **Independent** — occurrence of one does not affect the other
  $P(A \mid B) = P(A)$, and equivalently $P(A \cap B) = P(A)\,P(B)$

> **Note:** Nontrivial events cannot be both mutually exclusive **and** independent (unless a probability is 0).

---

2) Union, Conditional, and Product Rules

- **Additive rule** (mutually exclusive)
  $P(A \cup B) = P(A) + P(B)$
- **General union**
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- **Multiplicative rule**
  $P(A \cap B) = P(B)\,P(A \mid B) = P(A)\,P(B \mid A)$
- **Independence test**
  $A \perp B \iff P(A \cap B) = P(A)P(B) \iff P(A \mid B) = P(A)$

---

3) Screening / Diagnostic Testing (2×2 Table)

Let D = disease, T^+ = positive test.

- **Sensitivity** (true positive rate)
  $P(T^+ \mid D)$
- **Specificity** (true negative rate)
  $P(T^- \mid D^c)$
- **False rates**
  $\mathrm{FPR} = P(T^+ \mid D^c) = 1 - \text{specificity}$,
  $\mathrm{FNR} = P(T^- \mid D) = 1 - \text{sensitivity}$
- **Predictive values (representative samples)**
  $PPV = P(D \mid T^+), \qquad NPV = P(D^c \mid T^-)$
- **Bayes (expressed via sens/spec/prevalence)**

$$PPV = \frac{\text{sens} \cdot \text{prev}}{\text{sens} \cdot \text{prev} + (1 - \text{spec}) \cdot (1 - \text{prev})}$$

$$NPV = \frac{\text{spec} \cdot (1 - \text{prev})}{\text{spec} \cdot (1 - \text{prev}) + (1 - \text{sens}) \cdot \text{prev}}$$

- **Likelihood Ratios (quick Bayes updates)**
  $LR^+ = \dfrac{\text{sens}}{1-\text{spec}}, \quad LR^- = \dfrac{1-\text{sens}}{\text{spec}}$

---

4) Risk, Odds, RR, and OR

- **Relative Risk**
  $RR = \dfrac{P(D \mid E)}{P(D \mid E^c)}$
- **Odds Ratio** (probability form)
  $OR = \dfrac{P(D \mid E)/P(D^c \mid E)}{P(D \mid E^c)/P(D^c \mid E^c)}$
- **Rare disease link**
  $OR \approx RR$ when $P(D)$ is small in both groups

---

5) Descriptives: Quartiles, IQR, Variance, SD, CV

- **Quartile position** (for ordered data, sample size n)
  Q_p at position $n \cdot p/100$ (make sure to find next highest value in dataset)
- **Interquartile Range**
  $IQR = Q_3 - Q_1$
- **Population variance / SD**

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2, \qquad \sigma = \sqrt{\sigma^2}$$

- **Sample variance / SD**

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2, \qquad s = \sqrt{s^2}$$

- **Coefficient of Variation**
  $CV_{\text{pop}} = \dfrac{\sigma}{\mu} \times 100\%, \qquad CV_{\text{samp}} = \dfrac{s}{\bar{x}} \times 100\%$