Fundamentals

Sample Mean: $\overline{x} = x_1 + x_2 + \cdots + x_n = \frac{1}{n}\sum_{i=1}^{n} x_i$

Sample Median: Sort values in increasing order, then:

$$\tilde{x} = \begin{cases} \text{Middle value} & \text{If n is odd} \\ \text{Average of two middle values} & \text{If n is even} \end{cases}$$

Sample Range: $\text{Range}(x) = \text{Max}(x) - \text{Min}(x)$

Population mean: $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$

Population median: $\tilde{\mu} = \text{median of } \{x_1, x_2, \cdots, x_n\}$

Sample Variance: S^2 or $\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$

Standard Deviation: $\text{sd}(x) = \sqrt{\sigma^2} = \sigma$

Factorial: $!n = n \times (n-1) \times (n-2) \times \cdots \times 1$

Stem and Leaf Plots:

→ Each "stem" refers to the highest digits and each "leaf" is the lowest digit. This is the stem-and-leaf plot for:

{2 2 2 3 9 14 18 19 20 21 21 22 22 29 30 32 32 112}

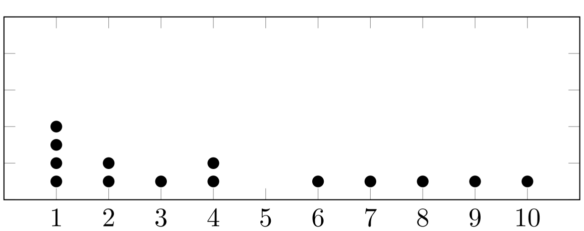
Stem	Leaves
0	2 2 2 3 9
1	4 8 9
2	0 1 1 2 2 9
3	0 2 2 3
4	4 7
5	
⋮	
11	2

Dot Plots:

→ Each "dot" over the x-axis references a single instance of that x-value in the set, this is the dot plot for:

{7 8 1 3 4 10 1 2 2 1 1 4 4 9 6}

Dot Plot



Box Plots:

→ Q_1 = First Quartile = median of the smallest half of values

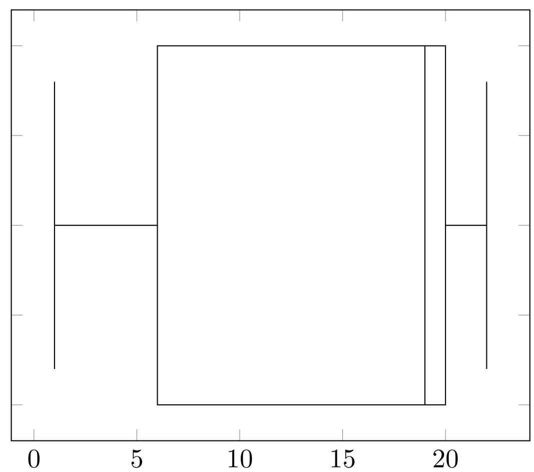
→ Q_3 = Second Quartile = median of largest half of values

→ $IQR = f_s$ = Fourth Spread = $Q_3 - Q_1$

Here's the recipe for constructing a boxplot ("cat and whisker plot"):

1. Draw a horizontal line that extends from the smallest to largest values in your data set.
2. Draw a rectangle with vertical lines at Q_1, Q_2, and Q_3. (Q_2 = median)
3. If $x_i Q_1 - 1.5 * IQR$ or $x_i > Q_3 + 1.5 * IQR$, then x_i is considered an outlier. Put a dot at the locations of outliers.
4. Draw whiskers that extend from the rectangle to the most extreme non-outlying observation.

Example: {1,5,7,18,20,22,50}, use dots instead of bars



Probabilities

Events: A, an event or set of events

Subset: $n \subset A$, a set made of elements in A

Probability: $P(A)$, the probability of event A occuring

→ $\sum_x P(x) = 1$

Cardinality: $\#A$, the number of elements in set A

Combination: $C(\#A, n)/\binom{\#A}{n} = \dfrac{\#A!}{n!(\#A-n)!}$

Permutation: $P(\#A, n) = \dfrac{\#A}{(\#A-n)!}$

Given: $P(A|B)$, the probability of A given B

$\rightarrow P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Independent if: $P(A \cap B) = P(A) \times P(B)$

Mutually Exclusive if: $P(A \cap B) = \emptyset$

Bayes' Theorem: $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$

Distributions

Expected Value: $\mathbb{E}(X)$, the sum of the value $\times$ probability of each element in a distribution

Probability Mass Function: $p(x) = P(X = x)$

$$\rightarrow p(x = X) = \begin{cases} p(x_1) & \text{if conditional} \\ p(x_2) & \text{if conditional} \\ \vdots & \\ p(x_n) & \text{if conditional} \end{cases}$$

$\rightarrow \sum_x p(x) = 1$

Cumulative Distribution Function: $F_X = P(X \leq x)$

$\rightarrow P(a < X \leq b) = F_X(b) - F_X(a)$

$\rightarrow 0 \leq \frac{d}{dx}[F'_X]$

$$\rightarrow F_X(x) = \begin{cases} a & 0 \\ x & f(x) \\ \vdots & \\ b & 1 \end{cases}$$

$\rightarrow \lim_{x \to \infty} F_X(x) = 1$

$\rightarrow \lim_{x \to -\infty} F_X(x) = 0$

Discrete Random Variables:

$\rightarrow \mathbb{E}(X)/\mu = \sum_x xp(x)$

$\rightarrow \mathbb{E}(X^2) = \sum_x x^2 p(x)$

$\rightarrow \text{var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$

$\rightarrow \text{sd}(X) = \sqrt{\text{var}(X)}$

Binomials: $X \sim \text{Binomial}(A, p)$

$\rightarrow \mathbb{E}(X) = n \times p$

$\rightarrow \text{var}(X) = n \times p(1-p)$

$\rightarrow \text{sd}(X) = \sqrt{\text{var}(X)}$

Poisson:

$\rightarrow \mathbb{E}(X) = \lambda = \text{var}(X)$

$\rightarrow \text{PMF}/P(X = k) = \dfrac{\lambda^k e^{-\lambda}}{k!}$

$\rightarrow P(k \text{ events in } t \text{ interval}) = \dfrac{(rt)^k e^{-rt}}{k!}$

Continuous Random Variables:

$\rightarrow \mathbb{E}(X) = \int_{-\infty}^{\infty} [xf(x)]\, dx$

$\rightarrow \text{var}(X) = \int_{-\infty}^{\infty} \left[(x - \mu_X)^2\right] f(x) dx$

$\rightarrow \text{sd}(X) = \sqrt{\text{var}(X)}$