Lecture 3
Measures of Dispersion

Prof Dr Saad Abed Madhi

Measures of Dispersion in Radiology

Dispersion describes how **spread out** data values are around a central point (mean or median).

In radiology, measuring spread helps in assessing variability in imaging findings, diagnostic test accuracy, and research outcomes.

1. Range

- **Definition:** Difference between the largest and smallest value.

- **Formula:**

$$\text{Range} = X_{max} - X_{min}$$

Example (CT lesion sizes in cm): 2.1, 3.0, 2.8, 4.2, 3.6

$$\text{Range} = 4.2 - 2.1 = \textbf{2.1 cm}$$

2. Mean Deviation (MD)

- **Definition:** Average of absolute deviations from the mean or median.

- **Formula:**

$$MD = \frac{\sum \left| X_i - \bar{X} \right|}{n}$$

- **Example (MRI tumor diameters in cm):** 2.0, 2.5, 3.0, 3.5, 4.0

  - Mean $= 3.0$

  - Deviations $= (1.0, 0.5, 0, 0.5, 1.0)$

  - $MD = 3.0 \div 5 = $ **0.6 cm**

3. Variance

- **Definition:** Average of squared deviations from the mean.

- **Formula:**

$$Variance = \frac{\sum (X_i - \bar{X})^2}{n}$$

Example: MRI tumor diameters 2.5, 3.0, 2.8, 3.2 → Variance = **0.067**

4. Standard Deviation (SD)

- **Definition:** Square root of variance; shows average deviation in same units as data.

- **Formula:**

$$SD = \sqrt{Variance}$$

- **Example:** Variance = 0.067 → SD = **0.26 cm**

5. Coefficient of Variation (CV)

- **Definition:** Relative variability expressed as % of mean.

- **Formula:**

$$CV = \frac{SD}{\bar{X}} \times 100\%$$

- **Example:** Mean lesion size = 3.0 cm, SD = 0.6 cm
  $\rightarrow$ CV = **20%**

6. Interquartile Range (IQR)

- **Definition:** Spread of middle 50% of values.
- **Formula:**

$$IQR = Q_3 - Q_1$$

- **Example (PET SUV):** 4.1, 4.5, 5.0, 5.6, 6.0, 6.5, 7.2
  - Q1 = 4.5, Q3 = 6.5
    → IQR = **2.0**

7. Finding Outliers

- Outliers are values that lie far from the rest of the data — may represent true clinical abnormalities or measurement errors.

Methods

- **(a) IQR Method:**
- Compute IQR = Q3 – Q1
- Outlier if:

$$X < Q_1 - 1.5(IQR) \quad \text{or} \quad X > Q_3 + 1.5(IQR)$$

- **Radiology Example (PET SUV values):**

  4.1, 4.5, 5.0, 5.6, 6.0, 6.5, 12.0

- Q1 = 4.5, Q3 = 6.5 → IQR = 2.0
- Upper cutoff = 6.5 + (1.5×2.0) = 9.5
- SUV = 12.0 is an **outlier** → may suggest a highly suspicious lesion.

(b) Z-Score Method:

- Compute **Z**-score:

$$Z = \frac{X - \bar{X}}{SD}$$

- If $|\mathbf{Z}| > 3 \rightarrow$ Outlier.

Radiology Example (Lesion size in cm):

Mean = 3.0, SD = 0.5

- A lesion of 5.0 cm $\rightarrow$

Z=(5.0−3.0)/0.5=4.0 $\rightarrow$ **Outlier**

Why Outlier Detection Matters in Radiology?

- Identifies **extreme imaging values** that may represent rare or aggressive disease.

- Helps detect **measurement/recording errors** in radiology studies.

- Ensures correct **statistical analysis**, since outliers can distort mean, SD, and hypothesis testing results.

Comparison of Dispersion Measures in Radiology

Measure	Formula	Radiology Example	Interpretation
Range	$X\text{max} - X\text{min}$	CT lesion sizes: 2.1 – 4.2 → Range = **2.1 cm**	Spread between smallest and largest lesion size.
Mean Deviation (MD)	$MD = \dfrac{\sum |X_i - \bar{X}|}{n}$		
Variance	$Variance = \dfrac{\sum (X_i - \bar{X})^2}{n}$	MRI tumors (2.5, 3.0, 2.8, 3.2) → Var = **0.067**	Quantifies squared spread around the mean.
Standard Deviation (SD)	$SD = \sqrt{Variance}$	Above variance → SD = **0.26 cm**	Tumors deviate ~0.26 cm from the mean size.
Coefficient of Variation (CV)	$CV = \dfrac{SD}{\bar{X}} \times 100\%$	Mean = 3.0 cm, SD = 0.6 → CV = **20%**	Variability relative to mean; useful for comparing modalities.
Interquartile Range (IQR)	$IQR = Q_3 - Q_1$	PET SUV: Q1=4.5, Q3=6.5 → IQR = **2.0**	Spread of middle 50% of SUV values.
Outliers (IQR method)	Outlier if $X < Q_1 - 1.5(IQR)$ or $X > Q_3 + 1.5(IQR)$	PET SUV values: 12.0 > 9.5 → **Outlier**	Identifies extreme imaging values (possible abnormal lesion).
Outliers (Z-score method)	$Z = \dfrac{X - \bar{X}}{SD}$		

Exercises on Dispersion Measures

1- The sizes of 5 brain lesions measured on MRI are: **2.4, 3.1, 2.9, 3.7, 4.2 cm**. Find the **range**.

2- Ultrasound measurements of gallstone diameter (cm): **1.0, 1.5, 2.0, 2.5, 3.0**. Find the **mean deviation about the mean**.

3- Tumor diameters on CT (cm): **3.0, 3.2, 3.5, 3.3**. Find the **variance**.

4- Using Exercise 3 data, calculate the **SD**.

5-PET SUV values of a tumor are: **5.0, 5.5, 6.0, 6.5, 7.0**. Find the **CV**.

6-SUV values: **3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0**. Find the **IQR**.

7-PET SUV values: **4.0, 4.5, 5.0, 5.5, 6.0, 12.0**. Identify outliers.

8- Lesion sizes: mean = 2.8 cm, SD = 0.4 cm. A lesion measures **4.0 cm**. Is it an outlier?

9- CT lesion sizes: Mean = 5.0 cm, SD = 1.0 cm MRI lesion sizes: Mean = 2.5 cm, SD = 0.5 cm Which has greater **relative variability**?

10- Chest X-ray lesion sizes (cm): **1.2, 1.5, 2.0, 2.2, 2.6**. Find **Range, Mean, Variance, SD, MD**.

MCQ on **measures of dispersion**

- **1. The simplest measure of dispersion is:**

- A. Mean deviation
  B. Range
  C. Variance
  D. Standard deviation
  E. Interquartile range

- **2. In a dataset of lesion diameters (cm): 2.0, 2.5, 3.0, 3.5, 4.0, the range is:**

- A. 1.5
  B. 2.0
  C. 2.5
  D. 3.0
  E. 4.0

- **3. The mean deviation is based on:**

- A. Squared deviations from the mean
  B. Absolute deviations from the mean
  C. Cubic deviations from the mean
  D. Ratio of deviations
  E. Square roots of deviations

- **4. If the variance of a dataset is 16, the standard deviation is:**

- A. 2
  B. 4
  C. 6
  D. 8
  E. 16

- **5. Which measure of dispersion is expressed in percentage form?**
- A. Range
  B. Variance
  C. Coefficient of variation
  D. Mean deviation
  E. Interquartile range

- **6. If the mean of tumor sizes is 5 cm and the standard deviation is 0.5 cm, the coefficient of variation is:**
- A. 5%
  B. 10%
  C. 15%
  D. 20%
  E. 25%

- **7. The variance is defined as:**
- A. Mean of deviations
  B. Mean of absolute deviations
  C. Mean of squared deviations from the mean
  D. Square root of mean deviation
  E. Range divided by 2

- **8. Which of the following measures is least affected by extreme values?**

- A. Range
  B. Mean deviation
  C. Standard deviation
  D. Interquartile range
  E. Variance

- **9. The interquartile range represents the spread of:**

- A. 25% of the data
  B. 50% of the data
  C. 75% of the data
  D. 100% of the data
  E. None of these

- **10. A standard deviation of zero indicates:**

- A. High variability
  B. Low variability
  C. All values are equal
  D. Presence of outliers
  E. Non-numeric data

- **11. If mean = 50 and SD = 0, what can be said about the data?**

- A. It is highly variable
  B. It has one unique value repeated
  C. It has equal frequencies
  D. It is skewed
  E. It contains outliers

- **12. In a normal distribution, about 68% of observations lie within:**
- A. ±1 SD
  B. ±2 SD
  C. ±3 SD
  D. ±4 SD
  E. ±5 SD
- **13. The square of the standard deviation is known as:**
- A. Range
  B. Variance
  C. Mean deviation
  D. IQR
  E. None of these
- **14. Outliers are best detected using:**
- A. Mean
  B. Mode
  C. Z-score or IQR methods
  D. Median
  E. Range

- **15. If Q1 = 3, Q3 = 7, then the IQR =**
- A. 3
  B. 4
  C. 5
  D. 6
  E. 7
- **16. A Z-score greater than +3 or less than −3 indicates:**
- A. No variability
  B. Mild deviation
  C. Possible outlier
  D. Normal variation
  E. Incorrect data entry
- **17. Which measure of dispersion has the same unit as the data?**
- A. Variance
  B. Standard deviation
  C. Squared deviation
  D. Mean square
  E. Range ratio

- **18. When comparing variability between datasets with different means, it is best to use:**
- A. Range
  B. Variance
  C. Standard deviation
  D. Coefficient of variation
  E. Mean deviation
- **19. If the IQR = 2.0, the lower quartile = 4.0, and the upper limit for outliers is Q3 + 1.5×IQR, then:**
- A. 5.0
  B. 6.0
  C. 7.0
  D. 8.0
  E. 9.0
- **20. Which of the following best describes the importance of dispersion measures in radiology?**
- A. Only shows the central value
  B. Quantifies variability in image-derived data
  C. Used for patient demographics only
  D. Measures image resolution
  E. Detects artifacts