

Dispersion, skewness and the normal distribution

Free study notes · 2 notes · 25 Oct 2026

Free study material from Medha AI. For notes, practice, mock tests and previous papers for your exam, start free at **app.medhaai.in**.

1. Dispersion: range, quartile deviation, mean deviation, standard deviation, variance and coefficient of variation
2. Moments, skewness (Karl Pearson and Bowley), kurtosis and the normal distribution

1 Dispersion: range, quartile deviation, mean deviation, standard deviation, variance and coefficient of variation

Dispersion measures how far the values of a series are scattered around their average: range and quartile deviation use positions, mean deviation and standard deviation use every value, and the coefficient of variation compares spread across series.

Core facts

- **Dispersion** = the degree of scatter of values about a central value. Two series can have the same mean and very different spread.
- **Absolute measures** are in the units of the data: range, quartile deviation (QD), mean deviation (MD), standard deviation (SD).
- **Relative measures** are pure numbers (coefficients), used to compare series in different units or with different averages.

Measure	Formula	Relative measure
Range	$R = L - S$	$(L - S)/(L + S)$
Quartile deviation	$(Q3 - Q1)/2$	$(Q3 - Q1)/(Q3 + Q1)$
Mean deviation	Average absolute deviation from A	MD/A (A = mean or median)
Standard deviation	$\sqrt{(\sum f(x - \bar{x})^2)/N}$	$CV = (\sigma/\bar{x}) \times 100$

- **Interquartile range** = $Q3 - Q1$; **QD** is half of it, so it is also called the **semi-interquartile range**.
- **Mean deviation** can be taken about the mean, median or mode; it is **least when taken about the median**.
- **Standard deviation** (σ) was introduced by **Karl Pearson** (1893); it is the **root-mean-square deviation from the mean**.
- **Variance** = σ^2 . Shortcut: $\sigma^2 = \sum x^2/n - \bar{x}^2$ (individual) or $\sum fx^2/N - \bar{x}^2$ (frequency).
- **Coefficient of variation** (Karl Pearson): **$CV = \sigma/\bar{x} \times 100$** . Lower CV = **more consistent / more uniform / more stable**.
- **Effect of origin and scale**: adding or subtracting a constant leaves range, QD, MD and SD unchanged; multiplying by k multiplies them by |k| and the variance by k^2 .
- **SD of the first n natural numbers** = $\sqrt{((n^2 - 1)/12)}$.
- **Lorenz curve** is a **graphic** measure of dispersion (of income or wealth); the line of equal distribution is the 45° diagonal.

Explained

1. Range. Range = largest – smallest value. For a grouped table, take the upper limit of the highest class minus the lower limit of the lowest class. It is quick, but depends only on the two extremes, so one outlier

changes it. Used in **quality control** charts, stock price and temperature reports.

Example 1. For 5, 8, 12, 15, 20 find the range and its coefficient. *Solution.* $R = 20 - 5 = 15$. Coefficient = $15/(20 + 5) = 15/25$. **Answer: 15; 0.6**

2. Quartile deviation. It ignores the lowest 25% and highest 25% of values, so it is not affected by extreme values and can be found for **open-end** classes. In individual and discrete series, Q_1 = size of the **$(n + 1)/4$ th** item and Q_3 = size of the **$3(n + 1)/4$ th** item after arranging in order.

Example 2. Find QD and its coefficient for 20, 28, 40, 12, 30, 15, 50, 33, 25, 45, 18. *Solution.* In order: 12, 15, 18, 20, 25, 28, 30, 33, 40, 45, 50; $n = 11$. $Q_1 = (11 + 1)/4 = 3$ rd item = 18. $Q_3 = 3 \times 12/4 = 9$ th item = 40. $QD = (40 - 18)/2 = 11$. Coefficient = $(40 - 18)/(40 + 18) = 22/58$. **Answer: QD 11; coefficient ≈ 0.379**

For a continuous series, $Q_1 = L + ((N/4 - cf)/f) \times h$ and $Q_3 = L + ((3N/4 - cf)/f) \times h$, where L is the lower limit of the quartile class, cf the cumulative frequency before it, f its frequency and h its width.

Example 3. Find QD and its coefficient.

Marks	0–10	10–20	20–30	30–40	40–50
Students	5	8	12	10	5

Solution. $N = 40$; cumulative frequencies 5, 13, 25, 35, 40. Q_1 : $N/4 = 10$ th item lies in 10–20. $Q_1 = 10 + ((10 - 5)/8) \times 10 = 16.25$. Q_3 : $3N/4 = 30$ th item lies in 30–40. $Q_3 = 30 + ((30 - 25)/10) \times 10 = 35$. $QD = (35 - 16.25)/2 = 9.375$. Coefficient = $18.75/51.25$. **Answer: QD 9.375; coefficient ≈ 0.366**

3. Mean deviation. The average of the **absolute** deviations (signs ignored) from the mean or median. $MD = \Sigma|x - A|/n$, or $\Sigma f|x - A|/N$ for a frequency table (with mid-points for classes).

Example 4. For 5, 8, 12, 15, 20 find MD about the mean and its coefficient. *Solution.* $\bar{x} = 60/5 = 12$. $|x - 12| = 7, 4, 0, 3, 8$; total 22. $MD = 22/5 = 4.4$. Coefficient = $4.4/12$. **Answer: MD 4.4; coefficient ≈ 0.367**

Example 5. Find MD about the median.

x	2	4	6	8	10
f	2	4	8	4	2

Solution. $N = 20$, median = 6 (the table is symmetric). $\Sigma f|x - 6| = 2(4) + 4(2) + 8(0) + 4(2) + 2(4) = 32$. $MD = 32/20$. **Answer: 1.6**

4. Standard deviation and variance.

- Individual series: $\sigma = \sqrt{(\Sigma(x - \bar{x})^2/n)} = \sqrt{(\Sigma x^2/n - \bar{x}^2)}$.
- Assumed mean (short-cut) method:** with $d = x - A$, $\sigma = \sqrt{(\Sigma fd^2/N - (\Sigma fd/N)^2)}$.
- Step deviation method:** with $d' = (x - A)/h$, $\sigma = h \times \sqrt{(\Sigma fd'^2/N - (\Sigma fd'/N)^2)}$.

Example 6. For 5, 8, 12, 15, 20 find variance, SD and CV. *Solution.* $\bar{x} = 12$. Squared deviations 49, 16, 0, 9, 64; total 138. $\sigma^2 = 138/5 = 27.6$; $\sigma = \sqrt{27.6} \approx 5.254$. $CV = 5.254/12 \times 100$. **Answer: variance 27.6; SD ≈ 5.25 ; CV $\approx 43.8\%$**

Example 7. Find the mean, SD and CV for the marks table of Example 3 by the step deviation method.

Solution. Mid-points 5, 15, 25, 35, 45; $A = 25$, $h = 10$, $d' = -2, -1, 0, 1, 2$.

m	f	d'	fd'	fd' ²
5	5	-2	-10	20
15	8	-1	-8	8
25	12	0	0	0
35	10	1	10	10
45	5	2	10	20

$\Sigma fd' = 2$, $\Sigma fd'^2 = 58$, $N = 40$. $\bar{x} = 25 + (2/40) \times 10 = 25.5$. $\sigma = 10 \times \sqrt{(58/40 - (2/40)^2)} = 10 \times \sqrt{(1.45 - 0.0025)} = 10 \times \sqrt{1.4475} \approx 12.03$. $CV = 12.03/25.5 \times 100$. **Answer: mean 25.5; SD $\approx$ 12.03; CV $\approx$ 47.2%**

5. Coefficient of variation and consistency.

Example 8. Factory A pays an average daily wage of ₹300 with SD ₹15; factory B pays ₹500 with SD ₹20. Where are wages more uniform? *Solution.* $CV(A) = 15/300 \times 100 = 5\%$; $CV(B) = 20/500 \times 100 = 4\%$. **Answer: factory B** (lower CV), even though its SD is larger.

6. Properties of SD.

- It is the most widely used measure: rigidly defined, based on all values, and suited to algebra (combined SD, sampling theory).
- $\Sigma(x - \bar{x})^2$ is **least** when deviations are taken from the mean, so SD about the mean is the smallest root-mean-square deviation.
- Independent of change of origin, not of scale.
- $\sigma = 0$ only when all values are equal.

Deeper

1. Combined standard deviation of two groups. $\sigma_{12} = \sqrt{[(n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)) / (n_1 + n_2)]}$, where $d_1 = \bar{x}_1 - \bar{x}_{12}$ and $d_2 = \bar{x}_2 - \bar{x}_{12}$, and $\bar{x}_{12} = (n_1\bar{x}_1 + n_2\bar{x}_2)/(n_1 + n_2)$.

Example 9. Section A: 50 students, mean 60, SD 8. Section B: 50 students, mean 50, SD 6. Find the combined SD. *Solution.* $\bar{x}_{12} = (3,000 + 2,500)/100 = 55$; $d_1 = 5$, $d_2 = -5$. $\sigma_{12}^2 = [50(64 + 25) + 50(36 + 25)]/100 = (4,450 + 3,050)/100 = 75$. **Answer: $\sigma_{12} = \sqrt{75} \approx 8.66$**

2. Correcting a wrong observation.

Example 10. 20 observations have mean 40 and SD 5. One value, 43, was wrongly copied as 53. Find the correct mean and SD. *Solution.* Wrong $\Sigma x = 20 \times 40 = 800$; correct $\Sigma x = 800 - 53 + 43 = 790$; correct mean = 39.5. Wrong $\Sigma x^2 = n(\sigma^2 + \bar{x}^2) = 20(25 + 1,600) = 32,500$; correct $\Sigma x^2 = 32,500 - 2,809 + 1,849 = 31,540$. Correct variance = $31,540/20 - 39.5^2 = 1,577 - 1,560.25 = 16.75$. **Answer: mean 39.5; SD = $\sqrt{16.75} \approx 4.09$**

3. Mean deviation of a grouped table. For the marks table ($\bar{x} = 25.5$), $\Sigma f|m - 25.5| = 5(20.5) + 8(10.5) + 12(0.5) + 10(9.5) + 5(19.5) = 385$, so MD about the mean = $385/40 = 9.625$.

4. Useful inequalities and relations.

- For any data, **MD about the mean $\leq$ SD**, and **SD $\leq$ Range/2**.

- For a **normal** distribution, approximately **QD** = $(2/3)\sigma$ and **MD** = $(4/5)\sigma$, so **QD : MD : SD** $\approx 10 : 12 : 15$ (exact values 0.6745σ and 0.7979σ).
- Variance of a sample: these notes, like most degree texts, divide by n . The unbiased sample estimate divides by **$n - 1$** ; the question will say so when it is wanted.

5. Lorenz curve (M.O. Lorenz, 1905). Plot cumulative percentages of the variable (income) on the Y axis against cumulative percentages of the units (persons) on the X axis, both from 0 to 100. The diagonal is the **line of equal distribution**. The farther the curve sags below the diagonal, the greater the inequality. The **Gini coefficient** = area between the curve and the diagonal $\div$ whole area under the diagonal (0 = perfect equality, 1 = perfect inequality).

6. Merits and demerits.

Measure	Merit	Demerit
Range	Simplest, quick	Only two extremes
QD	Not hit by extremes; open-end classes	Ignores half the data
MD	Uses all values; easy meaning	Ignores signs; poor algebra
SD	Uses all values; algebraic	Pulled by extremes; harder

Common traps

✗ "Coefficient of range = $(L - S)/L$." ✓ It is $(L - S)/(L + S)$.

✗ "QD = $Q3 - Q1$." ✓ That is the interquartile range; **QD** = $(Q3 - Q1)/2$.

✗ "Mean deviation is least about the mean." ✓ MD is least about the **median**; the **sum of squared deviations** is least about the mean.

✗ "Higher SD always means less consistent." ✓ Compare **CV** when the means differ.

✗ "SD changes when 10 is added to every value." ✓ SD is independent of **origin**; it changes only with **scale**.

✗ "Multiplying by 3 triples the variance." ✓ Variance becomes **9 times**; SD triples.

✗ "Variance = $(\sum x)^2/n - x^2$." ✓ It is $\sum x^2/n - x^2$.

✗ "Step deviation: $\sigma = \sqrt{(\sum fd^2/N - (\sum fd'/N)^2)}$." ✓ Multiply by the class width: **$\sigma = h \times \sqrt{(\dots)}$** .

Quick revision

- Range = $L - S$; coefficient $(L - S)/(L + S)$.
- QD = $(Q3 - Q1)/2$; coefficient $(Q3 - Q1)/(Q3 + Q1)$.
- $Q1 = (n + 1)/4$ th item; $Q3 = 3(n + 1)/4$ th item.

- $MD = \sum |x - A|/n$; least about the median.
- $\sigma^2 = \sum x^2/n - \bar{x}^2$; step deviation $\sigma = h\sqrt{(\sum fd'^2/N - (\sum fd'/N)^2)}$.
- $CV = \sigma/\bar{x} \times 100$; lower CV = more consistent.
- Origin change: no effect; scale $\times k$: $SD \times |k|$, variance $\times k^2$.
- Combined $\sigma^2 = [n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)]/(n_1 + n_2)$.
- Normal: $QD \approx (2/3)\sigma$, $MD \approx (4/5)\sigma$.
- Lorenz curve: graphic measure; farther from diagonal = more inequality.

2 Moments, skewness (Karl Pearson and Bowley), kurtosis and the normal distribution

Skewness measures the lack of symmetry of a distribution, kurtosis measures how peaked it is compared with the normal curve, and moments give one family of formulas for both; the normal distribution is the symmetric, mesokurtic benchmark.

Core facts

- **Skewness** = lack of symmetry. In a symmetric distribution **mean = median = mode**.
- **Positive (right) skew**: long tail to the right; **mean > median > mode**.
- **Negative (left) skew**: long tail to the left; **mean < median < mode**.
- **Karl Pearson's coefficient of skewness**: $Sk = (\text{Mean} - \text{Mode})/\sigma$; using $\text{Mode} = 3 \text{ Median} - 2 \text{ Mean}$, $Sk = 3(\text{Mean} - \text{Median})/\sigma$. Limits of the median form: **-3 to +3**.
- **Bowley's coefficient** (quartile measure): $Sk_B = (Q3 + Q1 - 2 \text{ Median})/(Q3 - Q1)$; limits **-1 to +1**.
- **Kelly's coefficient** (percentile measure): $(P90 + P10 - 2 \text{ Median})/(P90 - P10)$.
- **Moments about the mean** (central moments): $\mu_r = \sum f(x - \bar{x})^r/N$. **$\mu_1 = 0$, $\mu_2 = \sigma^2$** .
- **$\beta_1 = \mu_3^2/\mu_2^3$** (skewness), **$\beta_2 = \mu_4/\mu_2^2$** (kurtosis); **$\gamma_1 = \sqrt{\beta_1}$** with the sign of μ_3 , **$\gamma_2 = \beta_2 - 3$** .
- **Kurtosis** = peakedness (Karl Pearson, 1905):
 - **$\beta_2 = 3$: mesokurtic** (normal);
 - **$\beta_2 > 3$: leptokurtic** (more peaked);
 - **$\beta_2 < 3$: platykurtic** (flat-topped).
- **Normal distribution**: bell-shaped, symmetric about μ , mean = median = mode = μ , $\beta_1 = 0$, $\beta_2 = 3$, total area 1.
- **Areas**: $\mu \pm \sigma$ covers **68.27%**, $\mu \pm 2\sigma$ **95.45%**, $\mu \pm 3\sigma$ **99.73%**; $\mu \pm 1.96\sigma$ covers 95% and $\mu \pm 2.58\sigma$ covers 99%.
- **Standard normal variate**: **$Z = (X - \mu)/\sigma$** , with mean 0 and SD 1.

Explained

1. Tests of skewness. A distribution is skewed if (a) mean, median and mode differ; (b) the quartiles are not equidistant from the median ($Q3 - Md \neq Md - Q1$); (c) the curve has a longer tail on one side; (d) the total frequencies on either side of the mode are not equal.

Empirical relation (Karl Pearson) for moderately skewed data: **Mode = 3 Median - 2 Mean**, that is, $\text{Mean} - \text{Mode} = 3(\text{Mean} - \text{Median})$.

2. Karl Pearson's coefficient.

Example 1. Mean 50, median 48, SD 10. Find Karl Pearson's coefficient of skewness. *Solution.* $Sk = 3(50 - 48)/10 = 6/10$. **Answer: +0.6 (positive skew)**

Example 2. Mean 45, mode 50, SD 8. *Solution.* $Sk = (45 - 50)/8$. **Answer: -0.625 (negative skew)**

3. Bowley's coefficient. Useful for open-end classes and when the mode is ill-defined, since it uses only Q1, median and Q3. It ignores the 25% of values at each end.

Example 3. Q1 = 20, median = 30, Q3 = 50. Find Bowley's coefficient. *Solution.* $Sk_B = (50 + 20 - 60)/(50 - 20) = 10/30$. **Answer: +0.33** (Q3 is farther from the median than Q1)

4. Both measures on a grouped table.

Marks	0–10	10–20	20–30	30–40	40–50
Students	5	8	12	10	5

From the dispersion note (same table): mean = 25.5, $\sigma \approx 12.03$, Q1 = 16.25, Q3 = 35.

Example 4. Find Karl Pearson's (mode form) and Bowley's coefficients. *Solution.*

- Cumulative frequencies 5, 13, 25, 35, 40; $N/2 = 20$ lies in 20–30. Median = $20 + ((20 - 13)/12) \times 10 \approx 25.83$.
- Modal class 20–30 ($f_1 = 12$, $f_0 = 8$, $f_2 = 10$). Mode = $20 + (12 - 8)/(2 \times 12 - 8 - 10) \times 10 = 20 + (4/6) \times 10 \approx 26.67$.
- Karl Pearson: $Sk = (25.5 - 26.67)/12.03 \approx -0.097$.
- Bowley: $Sk_B = (35 + 16.25 - 2 \times 25.83)/(35 - 16.25) = (51.25 - 51.67)/18.75 \approx -0.022$.

Answer: about -0.10 and -0.02; a very slight negative skew

5. Moments.

- Raw moments about an arbitrary origin A: $\mu'_r = \Sigma f(x - A)^r/N$. About zero they are written v_r ($v_1 = \text{mean}$).
- Converting raw to central moments ($\mu'_1 = x - A$):
 - $\mu_2 = \mu'_2 - \mu'^2_1$
 - $\mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1$
 - $\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1$
- First moment about the mean is always 0; second is the variance; third measures skewness; fourth measures kurtosis.

Example 5. The first four moments about 4 are 1, 4, 10 and 45. Find the mean, the central moments and comment on skewness and kurtosis. *Solution.* Mean = $4 + 1 = 5$. $\mu_2 = 4 - 1 = 3$. $\mu_3 = 10 - 3(4)(1) + 2(1) = 0$. $\mu_4 = 45 - 4(10)(1) + 6(4)(1) - 3(1) = 26$. $\beta_1 = 0$; $\beta_2 = 26/9 \approx 2.89$. **Answer: symmetric ($\beta_1 = 0$) and slightly platykurtic ($\beta_2 < 3$)**

6. The normal curve. Probability density $f(x) = (1/(\sigma\sqrt{2\pi})) e^{-(x - \mu)^2/(2\sigma^2)}$, $-\infty < x < \infty$. It was derived by **De Moivre** (1733) and developed by **Laplace** and **Gauss**, so it is also called the **Gaussian** curve. Properties:

- bell-shaped, unimodal, symmetric about $x = \mu$;
- mean = median = mode = μ ; quartiles at $\mu \pm 0.6745\sigma$;
- the curve never touches the x-axis (**asymptotic**);
- **points of inflexion at $\mu \pm \sigma$** ;
- maximum ordinate at $x = \mu$, equal to $1/(\sigma\sqrt{2\pi}) \approx 0.3989/\sigma$;
- odd central moments are 0; $\mu_2 = \sigma^2$, $\mu_4 = 3\sigma^4$, so $\beta_1 = 0$ and $\beta_2 = 3$;
- QD $\approx 0.6745\sigma$ (about $2\sigma/3$) and MD $\approx 0.7979\sigma$ (about $4\sigma/5$);

- determined completely by its two parameters, μ and σ .

Example 6. Marks are normal with mean 60 and SD 10. What share of students score between 50 and 70? Above 80? *Solution.* 50 and 70 are $\mu \pm \sigma$: **68.27%**. 80 is $\mu + 2\sigma$; area beyond $z = 2$ is $(1 - 0.9545)/2 \approx 0.0228$.

Answer: about 68.3%; about 2.3%

Deeper

1. Using the normal table. Tables give the area from 0 to z . Useful values: $z = 0.5 \rightarrow 0.1915$; $z = 1 \rightarrow 0.3413$; $z = 1.5 \rightarrow 0.4332$; $z = 1.96 \rightarrow 0.4750$; $z = 2 \rightarrow 0.4772$; $z = 2.58 \rightarrow 0.4951$; $z = 3 \rightarrow 0.49865$.

Example 7. Marks of 1,000 students are normal with mean 60 and SD 10. How many score above 75?

Solution. $z = (75 - 60)/10 = 1.5$. Area above = $0.5 - 0.4332 = 0.0668$. $0.0668 \times 1,000 \approx 66.8$. **Answer: about 67 students**

Example 8. In the same marks distribution, how many of 500 students score between 50 and 65? *Solution.* $z = -1$ and $z = 0.5$. Area = $0.3413 + 0.1915 = 0.5328$; $\times 500 \approx 266.4$. **Answer: about 266 students**

2. Moments of an individual series.

Example 9. For 2, 3, 7, 8, 10 find μ_2 , μ_3 , μ_4 , β_1 and β_2 . *Solution.* $\bar{x} = 30/5 = 6$; deviations -4, -3, 1, 2, 4. $\mu_2 = (16 + 9 + 1 + 4 + 16)/5 = 46/5 = 9.2$. $\mu_3 = (-64 - 27 + 1 + 8 + 64)/5 = -18/5 = -3.6$. $\mu_4 = (256 + 81 + 1 + 16 + 256)/5 = 610/5 = 122$. $\beta_1 = (-3.6)^2/9.2^3 = 12.96/778.688 \approx 0.0166$; $\beta_2 = 122/9.2^2 = 122/84.64 \approx 1.44$. **Answer: slight negative skew ($\mu_3 < 0$); platykurtic ($\beta_2 < 3$)**

3. Effect of origin and scale. Central moments do not change with origin; under $y = (x - A)/h$, $\mu_r(x) = h^r \mu_r(y)$. Hence **β_1 and β_2 are pure numbers, independent of both origin and scale.**

4. Sheppard's corrections (grouped data, equal width h , frequencies tapering to zero at both ends):

- corrected $\mu_2 = \mu_2 - h^2/12$;
- μ_3 needs no correction;
- corrected $\mu_4 = \mu_4 - (1/2)h^2\mu_2 + (7/240)h^4$.

5. Other facts.

- $\beta_1 \geq 0$ always (it is a square), so it cannot show the **direction** of skew; use y_1 or the sign of μ_3 .
- $\beta_2 \geq 1$ for every distribution, and $\beta_2 \geq \beta_1 + 1$.
- Percentile coefficient of kurtosis $\kappa = QD/(P_{90} - P_{10})$; $\kappa \approx \mathbf{0.263}$ for the normal curve.
- Normal approximation: the **binomial** tends to the normal when n is large and p is not near 0 or 1; the **Poisson** tends to the normal when its mean is large. The **central limit theorem** makes sample means approximately normal, which is why most large-sample tests use z .
- Uses of the normal curve: sampling theory and tests of significance, quality control limits ($\mu \pm 3\sigma$), grading and standardising scores.

Common traps

X "In positive skewness, mean < median < mode." ✓ Positive skew: **mean > median > mode**.

✗ "Bowley's coefficient lies between -3 and +3." ✓ Bowley's lies between **-1 and +1**; Karl Pearson's median form lies between -3 and +3.

✗ "Karl Pearson's coefficient = (Mean - Median)/ σ ." ✓ It is **(Mean - Mode)/ σ** or **3(Mean - Median)/ σ** .

✗ "Empirical relation: Mode = 3 Mean - 2 Median." ✓ **Mode = 3 Median - 2 Mean**.

✗ "Leptokurtic means $\beta_2 < 3$." ✓ Leptokurtic (peaked) is **$\beta_2 > 3$** ; platykurtic (flat) is $\beta_2 < 3$.

✗ "The first moment about the mean is the mean." ✓ It is **0**; the second central moment is the **variance**.

✗ " β_1 is negative for negatively skewed data." ✓ β_1 is a square and never negative; the **sign of μ_3 (or γ_1)** shows direction.

✗ "Points of inflexion of the normal curve are at $\mu \pm 2\sigma$." ✓ They are at **$\mu \pm \sigma$** .

✗ "The normal curve meets the x-axis at $\mu \pm 3\sigma$." ✓ It **never** touches the axis; 99.73% of the area lies within $\mu \pm 3\sigma$.

Quick revision

- Symmetric: mean = median = mode; positive skew: mean > median > mode.
- Karl Pearson $Sk = (\text{Mean} - \text{Mode})/\sigma = 3(\text{Mean} - \text{Median})/\sigma$.
- Bowley $Sk = (Q_3 + Q_1 - 2Md)/(Q_3 - Q_1)$, between -1 and +1.
- Mode = 3 Median - 2 Mean.
- $\mu_1 = 0$, $\mu_2 = \sigma^2$; $\beta_1 = \mu_3^2/\mu_2^3$, $\beta_2 = \mu_4/\mu_2^2$.
- $\beta_2 = 3$ meso, > 3 leptokurtic, < 3 platykurtic; $\gamma_2 = \beta_2 - 3$.
- $\mu_3 = \mu'^3 - 3\mu'\mu_2' + 2\mu_1'^3$.
- Normal: inflexion at $\mu \pm \sigma$; asymptotic; $\beta_1 = 0$, $\beta_2 = 3$.
- Areas: 68.27%, 95.45%, 99.73%; $1.96\sigma \rightarrow 95\%$, $2.58\sigma \rightarrow 99\%$.
- $Z = (X - \mu)/\sigma$; $QD \approx 2\sigma/3$, $MD \approx 4\sigma/5$.