

代数 1 · 单元 1 · 评估指南

一变量统计

家长或教师指南——观察要点

1 单元返回数据显示和统计，新的重点是解释和 标准差 作为变异性的度量。观察学生是否能够阅读这些数字的含义，而不仅仅是计算它们。学生可能因为错误的原因而正确，也可能因为正确的原因而错误——两者都很重要。

问题 1 — 两个班级的睡眠数据

- (a) B 类具有更大的可变性 — 与 A 中的 6 到 10 的值从 4 分布到 12 相比。强响应：以视觉方式记录范围或描述分布，而不是计算 SD。
- (b) 诊断要点：中位数相同不代表分布相同。较好回答：明确指出“中位数只说明中间的位置，不能说明数据有多分散。”
- 弱：“它们是相同的，因为中位数相等。”这正是本单元针对的误解。

问题 2 - 左偏草图

- (a) 尾部应向左（低侧）。常见错误：学生画右斜，因为“向左倾斜”在视觉上听起来像“向左倾斜”。
- (b) 平均值 < 中位数。平均值被拉向长尾。
- 强：使用均值对极值敏感的概念进行解释。少数低分拖累了它。

问题 3 - 极值的计算和影响

- (a) 平均值 = 17.375。 (b) 中位数 = 16.5。 (c) $Q1 = 14$, $Q3 = 20.5$, $IQR = 6.5$ 。
- (d) 添加 80 后：新均值 ≈ 24.33 ，新中位数 = 18。关于 7 的平均跳跃；中位数左右移动 1.5。
- (e) ★ 诊断要点。较好回答：“平均数用到每个数值，因此受极端值影响；中位数只取决于中间，所以一个大数字并不能改变它太多。”

问题 4 — 均值相同，SD 不同

- 集合 Y 具有更大的 SD (30-70 与 X 中的 48-52 的值分布)。
- 强：“SD 衡量的是价差，Y 的值与 50 相差甚远。”
- 弱：尝试计算 SD；或者猜测具有较大值的集合具有较大的 SD（错误推理）。

问题 5 - 平均值与中位数的争论

- 目标是提供细致入微的答案。强：给出 中位数 更好的例子（倾斜数据 - 收入、房价）以及平均值更好的地方（对称数据；当您需要总计时）。
- 弱：在没有示例的情况下同意或不同意，或者只记住“中位数 对异常值具有鲁棒性”而不表明他们理解原因。

问题 6 — 对于常数数据，Han 的 $SD = 2$

- 实际 SD 为 0（零变异性）。强烈回应：“每个值都等于平均值，因此所有偏差都是 0，因此 SD 必须是 0。”
- Han 可能做了什么：将 SD 与其他东西混淆（数据的值？计数？）。或者错误地应用了公式。

- ★这个问题测试学生是否在概念上理解 SD，而不仅仅是程序上。

问题 7 — 租金箱线图 + 离群值规则

- (a) $IQR = 1300 - 900 = 400$ 。上栅栏 $= 1300 + 1.5(400) = 1900$ 。\$5500 > 1900，所以是的，它是一个异常值。
 - (b) 删除\$5500：平均数明显下降；中位数变化很小或根本没有变化；IQR 基本不变。
 - (c) 使用中位数——附近有豪华异常值；中位数代表“典型”租金较好。
 - 强：显示 $1.5 \cdot IQR$ 计算并将(b) 连接到中位数 的鲁棒性。

问题 8 - 两个篮球运动员，相同的平均值

- (a) 球员A — 稳定，SD 低。在势均力敌的比赛中可靠。
- (b) 播放器B — 高SD；35 游戏更有可能来自具有这种可变性的玩家。
- (c) ★概念顶峰：单个统计数据隐藏了很多内容。完整分布（或至少中心和传播）讲述了真实的故事。
 - 注意那些选择玩家B 进行(a) 的学生，因为他们有35 点游戏——他们没有考虑可靠性。

问题 9 — 设计一个统计问题

- 强：问题是真正的统计问题（答案涉及变异性）、确定了特定群体、描述了样本方法、选择的显示方式与数据类型相匹配，并且中心/分布选择反映了预期的形状。
- 弱：问题只有一个正确答案（“学校的入学人数是多少？”）——这不是统计数据。
- ★这是该单元的建模顶点；整体地对待这个标题。

问题 10 — 对/错

- (a) F — 相同的均值并不意味着相同的形状（参见问题 1）。
- (b) F — SD 是距离的度量；始终 ≥ 0 。
- (c) T — 将每个值乘以常数k 将SD 按 $|k|$ 缩放。
- (d) F — 向每个值添加常数确实会更改 中位数（它会移动），但不会更改 IQR 或 SD。
- (d) 是关键检查：学生要区分平移后的中心位置（改变）和离散程度（不变）。