

第五章 统计表与统计图

第一节 统计表

一、统计表的编制原则和结构

(一) 统计表的编制原则 (5 条硬规则)

1 一表一意，重点突出

一张表只讲一件事。

- 只围绕一个中心内容 / 一个主题
 - 内容多 → 拆成多张表
 - 否则信息密度反而下降，读者抓不到重点
-

2 主谓分明，按“句子”来设计表

统计表不是“堆数字”，而是一组被拆开的完整句子。

👉 经典结构逻辑：

- 标题：定语（说明研究对象、时间、地点）
- 横标目（左侧）：主语
- 纵标目（上方）：谓语
- 数据格：句子的具体数值

👉 从左向右读，每一行就是一句完整陈述。

3 层次清楚，分组有逻辑

- 标目的排列要：
 - 有层级

- 有顺序
 - 便于比较
 - 分组标准要前后一致，不能“混搭逻辑”
-

4 格式规范，能简就简

- 不追求“好看”，追求清晰
 - 文字、线条、符号一律从简
 - 单位、符号、记法要统一、规范
-

(二) 统计表的基本结构（5个组成部分）

表 5-1 试验药与对照药治疗帕金森病临床疗效的有效率比较

组别	例数	有效例数	无效例数	有效率/%
对照组	103	17	86	16.50
试验组	105	74	31	70.48
合计	208	91	117	43.75

(李康,贺佳主编, 2024, p. 42) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

一张合格的统计表，必须包含这 5 个元素：

1 标题 (Title)

核心要求

一眼看懂：研究什么 + 在哪 + 何时

- 放在表格正上方居中
- 简明但信息完整
- 多张表时必须编号（如：表 5-1）

◆ 好标题的判断标准：
脱离正文，单看标题也知道表在干什么。

2 标目

标目 = 数据 “是什么意思” 。

两类标目

- **横标目** (左侧)：说明每一行是谁
- **纵标目** (上方)：说明每一列是什么指标

必要时可设：

- **总标目**：概括纵标目的共同含义

◆ 标目要求：

- 简短
 - 单位明确 (如 %、‰)
 - 符号规范
 - 符合期刊或作业要求
-

3 线条 (三线表是标准答案)

论文和正式报告的默认格式：**三线表**

- **顶线**：标题与表体分隔
- **中线**：标目与数据分隔
- **底线**：表体结束

🚫 禁忌：

- 坚线
- 斜线

- 过多装饰线

◆ 分层需要时：

- 可用短横线辅助
- 不破坏整体简洁性

表 5-2 试验药物与安慰剂治疗帕金森病的多中心临床试验结果

医院	对照组(安慰剂)			试验组		
	例数	有效例数	有效率/%	例数	有效例数	有效率/%
医院 1	36	6	16.67	36	27	75.00
医院 2	31	4	12.90	34	22	64.71
医院 3	36	7	19.44	35	25	71.43
合计	103	17	16.50	105	74	70.48

(李康,贺佳主编, 2024, p. 43) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

4 数字

数字书写规则

- 一律阿拉伯数字
- 位数对齐
- 小数位数统一

特殊情况处理

- 无数字：—
- 数据缺失：... (并在备注说明)
- 数值为零：必须写“0”
- 表内不留空格

👉 空格 = 信息不完整 = 不规范

5 备注 (Footnotes)

使用原则

- 表内不塞说明性文字
- 需要说明的地方：
 - 用 *、† 等符号标出
 - 统一写在表下方

常见说明内容：

- 数据来源
- 特殊定义
- 统计方法
- 异常处理说明

二、统计表的种类

1 简单表 (Simple Table)

一句话定义

只有一个分组标志的统计表。

结构特征

- 标目只有一个层次
- 主语只按一个标志排列
 - 通常放在**横标目**
- **纵标目**：统计指标名称

适用场景

- 只比较一个因素
 - 结构清晰、对比直接
 - 适合：
 - 教学
 - 初步描述
 - 单因素比较
-

读表逻辑 (非常重要)

从左向右读，每一行就是一句完整的话。

例如：

- “试验组的例数为 105 例，其中有效例数为 74 例，无效例数为 31 例，有效率为 70.48%。”

◆ 判断是否是合格简单表的标准：
每一行能不能被完整读成一句话。

2 复合表 (Compound Table)

一句话定义

有两个及以上分组标志的统计表。

结构特征

- 标目有多层次
- 主语按多个标志同时排列
- 常见组合：
 - 地区 × 处理组
 - 年龄 × 性别
 - 医院 × 用药方案

标目安排原则 (很关键)

- **主要分组标志** (分项多、最重要)
👉 放在**横标目**
- 其余分组标志
👉 放在**纵标目 / 总标目**

📌 原则一句话：

“**谁是分析重点，谁就放在横标目。**”

复合表的价值

- **一次表格，多维分析**
- 可以同时比较：
 - 不同医院之间
 - 不同处理方法之间
- 信息密度高，但**对结构要求也更高**

3 简单表 vs 复合表 (快速对照)

维度 简单表 复合表

分组标志 1 个 ≥ 2 个

标目层次 单层 多层

信息密度 低 高

阅读难度 低 较高

适用分析 单因素 多因素

三、编制统计表的注意事项

1 统计表的格式：不是固定模板，而是 “**为目的服务**”

核心原则一句话

表的格式必须服从“你想让读者比较什么”。

表 5-3 试验组和对照组血压的基线情况

分组	例数	舒张压/mmHg		收缩压/mmHg	
		$\bar{X} \pm S$	M	$\bar{X} \pm S$	M
试验组	61	98.64 ± 4.73	98.0	150.56 ± 13.27	149.0
对照组	62	98.31 ± 5.29	96.0	148.23 ± 12.43	146.0

(李康,贺佳主编, 2024, p. 43) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

(1) 默认逻辑(最常用、最安全)

- **左边:** 比较主体(分组)
- **右边:** 统计指标与数据

👉 这是你在**没有特殊理由时**的首选结构。

(2) 可以“破例”的情况

表 5-4 试验组和对照组血压的基线情况

指标/mmHg	试验组($n_1=61$)		对照组($n_2=62$)	
	$\bar{X} \pm S$	M	$\bar{X} \pm S$	M
舒张压	98.64 ± 4.73	98.0	98.31 ± 5.29	96.0
收缩压	150.56 ± 13.27	149.0	148.23 ± 12.43	146.0

(李康,贺佳主编, 2024, p. 44) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

当出现以下情形时，可以调整标目位置：

- 指标种类很多
- 横向放不下
- 重点不是“组别差异”，而是“指标水平本身”

👉 这时可以：

- 把指标放在左边
- 把组别放在上方

❖ 关键不是“规范感”，而是：

哪种结构更利于表达主要信息。

2 统计表的内容结构：宁拆表，不硬塞

一条硬红线

一个表里，只放“可比的东西”。

什么情况一定要拆表？

- 指标性质不同
 - 时间 × 年龄 × 职业
- 分组逻辑不一致
- 读者无法一眼看出“在比什么”

👉 结果就是：

- 内容繁杂
 - 逻辑混乱
 - 比较失效
-

经典错误 (例 5-1 的问题本质)

【例 5-1】某地某年 111 例钩端螺旋体病患者发病季节、年龄和职业构成资料如表 5-5 所示,请按照编制统计表的基本要求,检查此表编制是否合适,如不合适,加以修改。

表 5-5 流行病学有关的主要因素(原表)

季节							
8月上旬		8月中旬		8月下旬		9月上旬	
人数	%	人数	%	人数	%	人数	%
7	6.3	70	63.1	28	25.2	6	5.4
年龄/岁							
15 以下		15~		45 以上			
人数	%	人数	%	人数	%		
12	10.8	95	85.6	4	3.6		
职业							
农民		学生		待业青年		其他	
人数	%	人数	%	人数	%	人数	%
106	95.5	1	0.9	3	2.7	1	0.9

表 5-5 的缺点是:①标题不够明确,而且未注明时间、地点;②发病季节、年龄和职业三项指标无可比性,放在一起,内容繁杂;③标目设计不合理,“人数”“%”多次重复,不利于分析对比;④表内文字和线条过多。因此可以将其修改成表 5-6、表 5-7 和表 5-8 三个表。

(李康,贺佳主编, 2024, p. 44) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

原表的问题不在“数字”,而在**逻辑**:

1. 标题不清楚 (无时间、地点)
2. 季节 / 年龄 / 职业 **不可比**
3. “人数” “%” 反复出现, 结构冗余
4. 表内线条和文字过多, 干扰阅读

正确改法 (非常重要的套路)

一个分析维度 = 一张表

所以拆成：

- 发病季节分布表

表 5-6 某地某年钩端螺旋体病患者发病季节分布(修改表)

季节	人数	构成比/%
8月上旬	7	6.3
8月中旬	70	63.1
8月下旬	28	25.2
9月上旬	6	5.4
合计	111	100.0

(李康,贺佳主编, 2024, p. 44) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

- 发病年龄分布表

表 5-7 某地某年钩端螺旋体病患者发病年龄分布(修改表)

年龄/岁	人数	构成比/%
< 15	12	10.8
15~< 45	95	85.6
≥ 45	4	3.6
合计	111	100.0

(李康,贺佳主编, 2024, p. 44) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

- 发病职业分布表

表 5-8 某地某年钩端螺旋体病患者发病职业分布(修改表)

职业	人数	构成比/%
农民	106	95.5
学生	1	0.9
待业青年	3	2.7
其他	1	0.9
合计	111	100.0

(李康,贺佳主编, 2024, p. 45) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

👉 拆表不是“多此一举”，而是让信息重新变得可读、可比。

3 标目设计：少层次、少重复

操作级原则

- 标目层次 尽量 ≤ 3 层
- 能合并的就合并
- 能用“构成比/%”表达的，不反复写“人数 + %”

👉 重复出现“人数”“%”，说明：

- 表设计失败
 - 信息组织不合理
-

4 标题：不是装饰，是“表的说明书”

合格标题必须包含

- 研究对象
- 时间
- 地点
- 研究内容

◆ 【例 5-2】的错误本质：

【例 5-2】某医院使用某药物治疗急慢性肝炎患者 161 例,疗效资料如表 5-9 所示,指出其缺点并加以改进。

表 5-9 使用某药物治疗急慢性肝炎疗效观察(原表)

效果 总例数	有效						无效	
	小计		近期痊愈		好转			
	例	%	例	%	例	%	例	%
	108	67.1	70	43.5	38	23.6	53	32.9

此表的主要目的是表达使用某药物治疗急慢性肝炎的疗效情况,但是其标题过于简单,主谓安排不合理,标目组合重复。可进行如下修改,见表 5-10。

(李康,贺佳主编, 2024, p. 45) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

- 标题过于笼统
- 看不出“谁、何时、在哪、干了什么”

修改后的好标题特征

脱离正文，单看标题也能理解表的研究背景。

5 统计表不是“结果堆砌”，而是“分析工具”

例 5-2 的关键问题

- 主谓不清
 - 标目组合重复
 - 表面“完整”，实际**难以比较**
-

修改后的思路（必须学会）

- 明确目的：表达疗效构成
- 用：
 - 疗效类别（左）
 - 例数 + 构成比（右）
- 让读者：
 - 一眼看到疗效分布

表 5-10 某年某医院使用某药物治疗急慢性肝炎的疗效观察(修改表)

疗效	例数	构成比/%
无效	53	32.9
好转	38	23.6
近期痊愈	70	43.5
合计	161	100.0

(李康,贺佳主编, 2024, p. 45) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

👉 好表是“读者不用思考也不会读错”的表。

第二节 统计图

一、统计图的制作

(一) 统计图的制作原则

① 图是为“分析目的”服务的

先问一句：我想让读者看懂什么？

- 比大小？
- 看趋势？
- 看构成？
- 看相关性？

👉 图形类型必须由“资料性质 + 分析目的”决定

而不是“我觉得这个图好看”。

2 一图一意，不要贪多

- 一幅统计图：
 - 只表达一个中心内容
 - 只对应一个统计指标

👉 多个指标 → 多张图

否则只会：

- 信息拥挤
- 重点模糊
- 读者误解

3 准确优先于美观，但两者都要

- 点要准
- 比例要真
- 线条、颜色、符号要区分清楚
 - 实线 / 虚线 / 点线
 - 不同颜色但不过度

👉 好图的底线：不会“看起来对，其实错”。

(二) 统计图的结构 (5 个必备部件)

一幅合格统计图，缺一不可。

1 标题 (Title)

功能一句话

告诉读者：这张图在说什么、在哪、什么时候。

- 一般放在**图下方居中**
- 需编号，便于引用

👉 判断标准：

脱离正文，单看标题就能理解图的研究背景。

2 图域 (Plot area)

- 图形展示的空间
- 多数统计图：
 - 用**直角坐标系第一象限**
- 圆图是例外

👉 图域要：

- 不拥挤
 - 不留大量空白
 - 比例自然
-

3 标目 (Axis labels)

两个问题必须回答清楚：

- 横轴是什么？
- 纵轴是什么？

👉 要求：

- 含义明确
 - **单位必须标出**
 - 不写“值”“数据”这种废话
-

4 图例 (Legend)

作用

解释“颜色 / 线型 / 图案”代表什么。

- 通常放在：
 - 横标目与标题之间
- 图域空间够时：
 - 也可放在图内

👉 原则：

- 能不放就不放
 - 一旦放，就必须清楚、不遮挡
-

5 刻度 (Scale)

这是最容易被忽视、却最容易造假的地方。

- 数值：
 - 由小到大
- 方向：
 - 纵轴：下 → 上
 - 横轴：左 → 右
- 原点与间隔：
 - 必须合理
 - 与数据量级匹配

👉 刻度选错，图会“看起来差别很大”，但其实是视觉误导。

(三) 统计图的种类 (先建立“选图意识”)

常用基础图形 (你必须熟)

图形	主要用途
直条图	不同组间大小比较
百分条图	构成比例比较
圆图	构成结构展示
普通线图	随时间变化趋势
半对数线图	指数型增长 / 下降
直方图	连续型数据分布
箱式图	分布、离散、异常值
误差条图	均数 ± 变异
散点图	两变量相关性

常见专业图 (知道即可)

- 森林图 (多中心研究)
- 生存曲线 (生存分析)
- 热图、树形图 (聚类分析)

👉 原则一句话：

“图不是越高级越好，而是越贴合问题越好。”

二、描述定量数据的统计图

1 直方图 (Histogram) ——看 “分布形态”

表 5-11 广州市 150 名 3 岁女孩身高的频数分布

组段/cm	频数	组段/cm	频数	组段/cm	频数
80~< 82	1	90~< 92	23	100~< 102	6
82~< 84	3	92~< 94	26	102~< 104	2
84~< 86	8	94~< 96	24	104~106	1
86~< 88	10	96~< 98	17		
88~< 90	19	98~< 100	10		

(李康,贺佳主编, 2024, p. 46) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

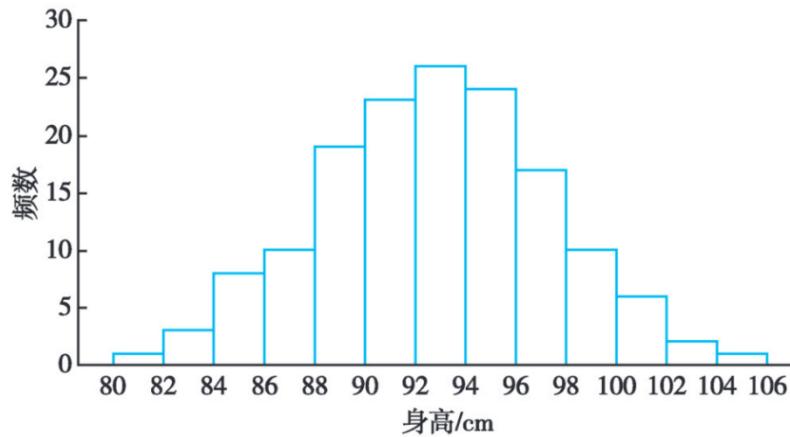


图 5-1 广州市 150 名 3 岁女孩身高频数分布

(李康,贺佳主编, 2024, p. 46) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

表 5-12 糖尿病患者和正常对照 HbA1c 分布

组段/%	频数		组段/%	频数		组段/%	频数	
	糖尿病	正常		糖尿病	正常		糖尿病	正常
4.0~< 5.2	1	20	6.8~< 7.2	7	2	8.8~< 9.2	4	0
5.2~< 5.6	2	28	7.2~< 7.6	16	2	9.2~< 9.6	8	0
5.6~< 6.0	3	27	7.6~< 8.0	12	1	9.6~< 10.0	5	0
6.0~< 6.4	3	13	8.0~< 8.4	10	1	10.0~12.8	19	0
6.4~< 6.8	7	6	8.4~< 8.8	3	0			

(李康,贺佳主编, 2024, p. 46) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

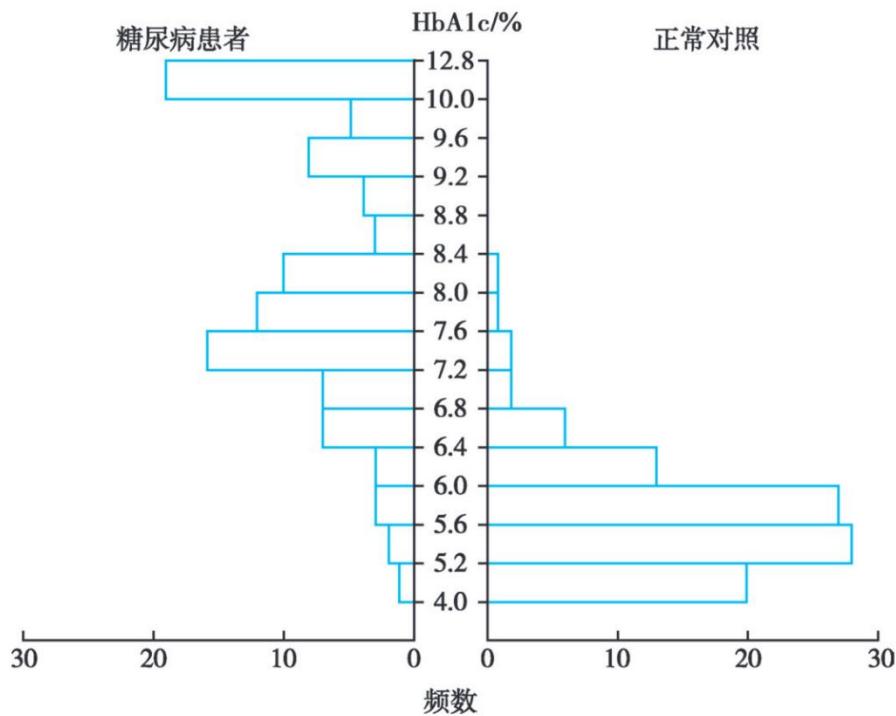


图 5-2 糖尿病患者和正常对照 HbA1c 的频数分布

(李康,贺佳主编, 2024, p. 47) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

核心用途

描述连续型变量的频数分布形态

典型问题：

- 数据集中在哪？
- 是否近似正态？
- 是否偏态？是否有长尾？

本质规则（必须记）

- 矩形面积 = 频数 / 频率
- 各矩形面积之和 = 总频数

坐标轴要求

- 纵轴必须从 0 开始
- 横轴：按变量实际范围分组
- 矩形：
 - 高度 = 频数（或频率）
 - 宽度 = 组距
- **组距不等时：**

$$\text{矩形高度} = \frac{\text{该组频数}}{\text{组距}}$$

◆ 直方图不是“条形图”，宽度有统计意义。

2 线图 (Line Chart) ——看“变化趋势”

核心用途

描述一个指标随另一个变量（通常是时间）的变化过程

表 5-13 我国 2001—2019 年肺癌和结直肠癌死亡率

单位:1/10 万

年份	肺癌		结直肠癌	
	死亡率	对数值	死亡率	对数值
2001	30.46	1.48	9.32	0.97
2002	31.94	1.50	10.01	1.00
2003	33.31	1.52	10.74	1.03
2004	35.03	1.54	11.60	1.06
2005	36.13	1.56	12.16	1.08
2006	36.60	1.56	12.47	1.10
2007	37.72	1.58	12.90	1.11
2008	39.14	1.59	13.41	1.13
2009	40.68	1.61	13.99	1.15
2010	42.28	1.63	14.61	1.16
2011	43.37	1.64	15.05	1.18
2012	44.13	1.64	15.31	1.18
2013	44.88	1.65	15.47	1.19
2014	45.94	1.66	15.72	1.20

(李康,贺佳主编, 2024, p. 47) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

续表

年份	肺癌		结直肠癌	
	死亡率	对数值	死亡率	对数值
2015	46.98	1.67	15.89	1.20
2016	48.24	1.68	16.28	1.21
2017	49.64	1.70	16.79	1.23
2018	51.28	1.71	17.55	1.24
2019	53.23	1.73	18.40	1.26

(李康,贺佳主编, 2024, p. 48) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

(1) 普通线图 (算术坐标)

- 横轴、纵轴都是算术刻度
- **强调绝对变化量**
- 适合回答：

“增了多少？”

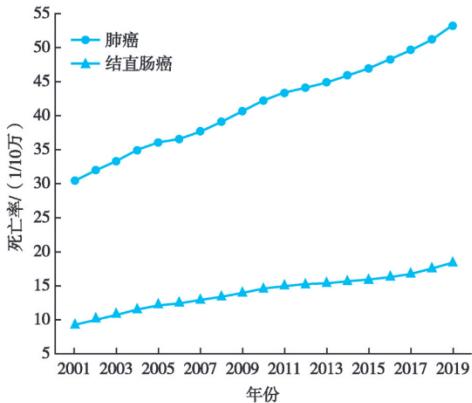


图 5-3 我国 2001—2019 年肺癌和结直肠癌死亡率的线图

(李康,贺佳主编, 2024, p. 48) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

(2) 半对数线图 (Semi-log)

- 纵轴：对数刻度
- 横轴：算术刻度
- **强调相对变化率 / 变化速度**
- 适合回答：

“涨得快不快？”

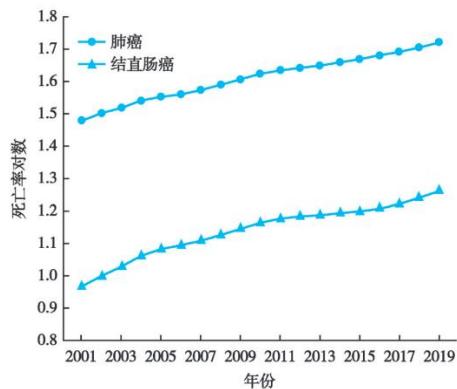


图 5-4 我国 2001—2019 年肺癌和结直肠癌死亡率的半对数线图

(李康,贺佳主编, 2024, p. 48) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

❖ 重要认知：

普通线图 “看幅度” ,
半对数线图 “看速度” 。

使用红线

- 折线 ≠ 光滑曲线
- 坐标轴**不一定**从 0 开始
- 纵横轴比例会影响视觉判断, 必须谨慎

3 箱式图 (Box Plot) ——看 “分布 + 离散 + 异常”

核心用途

比较多组数据的分布特征, 特别适合偏态数据

图形元素含义

- 中位数：箱内横线
- 箱体长度：四分位数间距 (IQR)
- 箱体越长 → 离散程度越大
- 箱体偏移中心 → 分布不对称

表 5-14 正常鼠和肝癌模型鼠血清中腺苷脱氨酶活性 单位: IU/ml

正常肝	肿瘤肝	正常肝	肿瘤肝
1.9	22.7	9.1	10.4
3.0	33.9	3.5	21.1
4.3	13.0	6.8	34.6
7.0	59.2	1.5	13.3
6.4	40.5	0.6	81.4

(李康,贺佳主编, 2024, p. 49) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

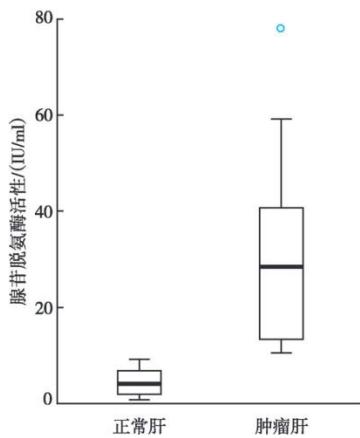


图 5-5 正常鼠和肝癌模型鼠血清中腺苷脱氨酶活性的箱式图

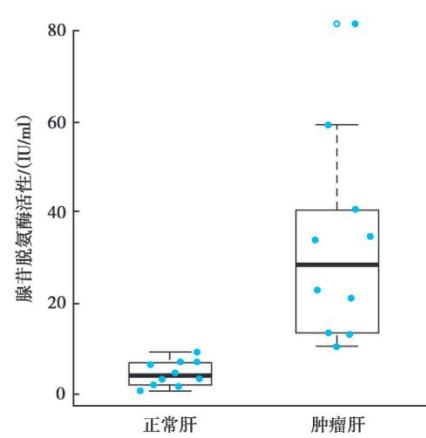


图 5-6 正常鼠和肝癌模型鼠血清中腺苷脱氨酶活性的箱式散点图

(李康,贺佳主编, 2024, p. 49) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

异常值判定 (必须会)

- 异常值:

$$< P_{25} - 1.5Q \text{ 或 } > P_{75} + 1.5Q$$

- 极端值：

$$< P_{25} - 3Q \text{ 或 } > P_{75} + 3Q$$

👉 箱式图纵轴**不要求从 0 开始**。

💡 误差条图 (Error Bar) ——看“均值 + 不确定性”

表 5-15 试验组和对照组患者 6 周后血压的变化情况

单位:mmHg

试验组		对照组		试验组		对照组	
舒张压	收缩压	舒张压	收缩压	舒张压	收缩压	舒张压	收缩压
12	17	10	10	8	20	7	10
12	8	-4	-7	-7	1	6	10
10	19	8	21	9	26	8	8
0	24	12	13	14	44	-4	6
16	9	10	-11	6	11	8	15
10	20	10	6	14	15	8	16
14	19	16	25	16	26	11	10
14	19	12	12	20	27	-4	3
11	14	10	7	9	10	11	25
2	8	2	12	2	9	9	5

(李康,贺佳主编, 2024, p. 50) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

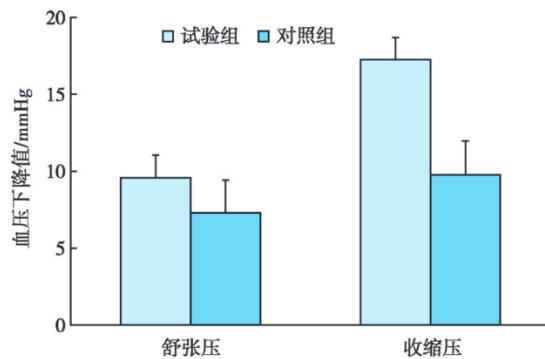


图 5-7 试验组和对照组血压的平均变化情况 ($\bar{X} \pm SE$)

(李康,贺佳主编, 2024, p. 49) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

核心用途

比较多组均值及其变异程度

常见表示

- 均值 \pm 标准差 (SD)
 - 均值 \pm 标准误 (SE)
 - 均值 \pm 置信区间 (CI)
-

使用红线 (非常重要)

- 纵轴必须从 0 开始
- 各组条宽必须一致
- 间距可不等
- 必须说明:
 - 用的是 SD / SE / CI

👉 不说明误差类型 = 图不合格。

5 散点图 (Scatter Plot) ——看“变量关系”

表 5-16 高血脂患者载脂蛋白 B (ApoB) 与低密度脂蛋白 (LDL) 的关系

ApoB/(g/L)	LDL/(mol/L)	Apo-B/(g/L)	LDL/(mol/L)
0.80	2.98	0.66	2.28
0.60	2.76	0.67	2.34
0.90	3.22	1.12	4.84
0.80	3.84	0.97	3.89
0.90	3.12	0.80	2.14
0.90	4.18	0.94	3.60
0.70	2.00	0.82	2.30
0.70	2.94	0.79	3.30
0.80	3.12	0.76	2.20
1.21	5.33	0.80	2.12

(李康, 贺佳主编, 2024, p. 50) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

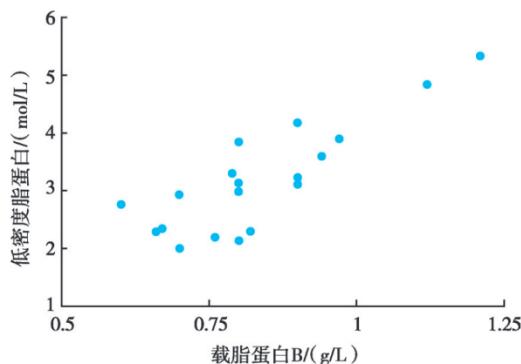


图 5-8 20 名高血脂患者载脂蛋白 B 与低密度脂蛋白关系的散点图

(李康, 贺佳主编, 2024, p. 50) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

核心用途

描述两个定量变量之间的关系

基本规则

- 一个点 = 一个观测值
- 横轴：自变量
- 纵轴：因变量
- 坐标轴起点不必从 0 开始

◆ 散点图只说明：

- 是否有关联
 - 趋势如何
- 不等于因果关系。**

6 热图 (Heat Map) ——看“模式与结构”

核心用途

- 多变量数据展示
- 聚类结果
- 表达量 / 强度分布

解读要点

- 颜色深浅 = 数值大小
 - 更强调：
 - 模式
 - 分群
 - 结构
- 而非精确数值

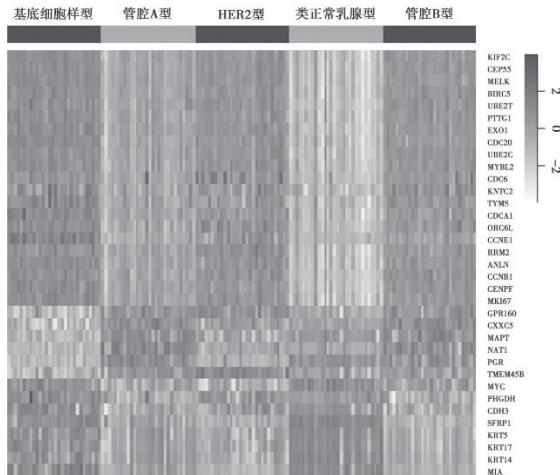


图 5-9 不同乳腺癌分型的 35 个基因表达热图

(李康,贺佳主编, 2024, p. 51) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

7 森林图 (Forest Plot) ——看“效应量 + 证据一致性”

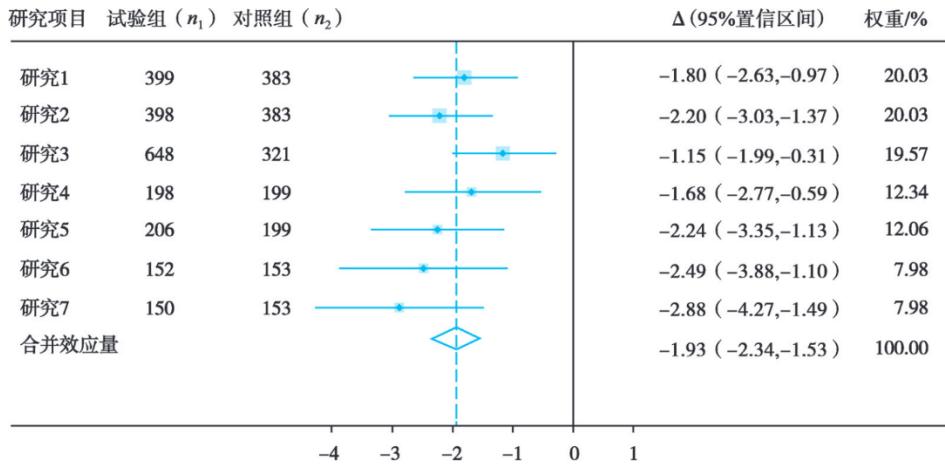


图 5-10 7 项以安慰剂为对照的多奈哌齐治疗阿尔茨海默病研究文献的综合分析

(李康,贺佳主编, 2024, p. 51) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

核心用途

综合多项研究或多中心结果

图形逻辑

- 每条横线：一个研究的效应量 + CI
- 垂直线：无效线
- 菱形：合并效应量

◆ 森林图是：

- **结果图**
 - 不是原始数据图
-

8 一张“选图决策速查表”

想回答的问题 首选图形

数据分布如何 直方图

随时间怎么变 线图

速度是否不同 半对数线图

分布/离散/异常 箱式图

均值差异 误差条图

两变量关系 散点图

多变量模式 热图

多研究合并 森林图

三、描述定性数据的统计图

(一) 直条图 (Bar chart) —— 用来“比大小”

一句话本质

直条图 = 比较相互独立指标的大小差异

- 表示的是：谁大、谁小、大多少

- 数据可以是：
 - 绝对数（人数、例数）
 - 相对数（率、比例）
-

① 单式条图

表 5-17 我国人口 1990 年和 2019 年四种主要死亡原因的死亡率 单位：1/10 万

死亡原因	1990 年	2019 年
脑卒中	116.34	153.91
慢性阻塞性肺疾病	105.09	72.94
缺血性心脏病	51.34	131.75
下呼吸道感染	45.95	13.03

(李康,贺佳主编, 2024, p. 52) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

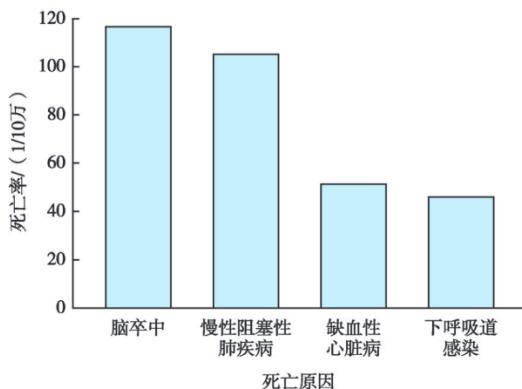


图 5-11 1990 年我国人口四种主要死亡原因的死亡率

(李康,贺佳主编, 2024, p. 52) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

结构特征

- 1 个统计指标
- 1 个分组标志（如：死因）

适用场景

- 单一时间点
- 单一分类
- 强调“排序与高低”

例子直觉理解：

1990年四种主要死因中，哪一种死亡率最高

2 复式条图

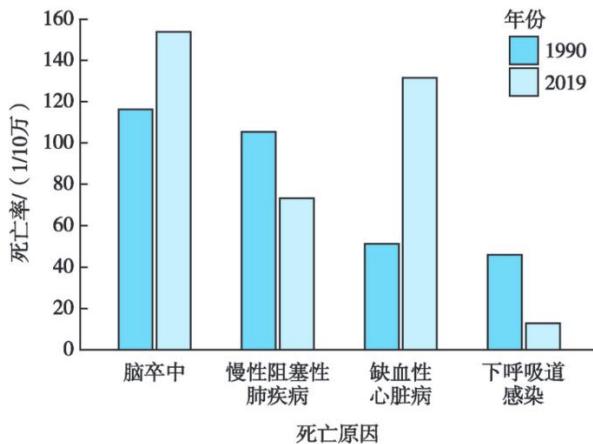


图 5-12 1990 年和 2019 年我国人口四种主要死亡原因的死亡率

(李康, 贺佳主编, 2024, p. 52) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

结构特征

- 1个统计指标
- ≥ 2 个分组标志（如：死因 \times 年份）

适用场景

- 同一指标在不同条件下的对比
- 强调“变化”和“差异”

核心优势：

把“同一类”放在一起比，而不是把“同一年”放在一起看

直条图的硬性规范（考试 + 实务都会卡）

- 横轴：分组
- 纵轴：数值
- 纵轴必须从 0 开始
- 刻度必须等距
- 所有直条：
 - 宽度相等
 - 间距等距
- 分组标志 不宜超过 3 组
 - 👉 超了就开始乱、开始误读

(二) 构成图——用来“看比例结构”

构成图回答的不是“谁多”，而是：

“这一整体里，各部分各占多少”

只适用于：构成比数据。

表 5-18 三个不同地区血型样本的频数分布和构成比(%)

地区	例数	A 型	B 型	AB 型	O 型
亚洲	1 080	321 (29.7)	369 (34.2)	95 (8.8)	295 (27.3)
欧洲	517	258 (49.9)	43 (8.3)	22 (4.3)	194 (37.5)

(李康,贺佳主编, 2024, p. 52) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

续表

地区	例数	A型	B型	AB型	O型
北美洲	995	408(41.0)	106(10.7)	37(3.7)	444(44.6)
合计	2 592	987(38.1)	518(20.0)	154(5.9)	933(36.0)

(李康,贺佳主编, 2024, p. 53) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

① 圆图 (Pie chart) ——看“单一整体的内部结构”

一句话本质

圆 = 100%，扇形 = 各部分所占比例

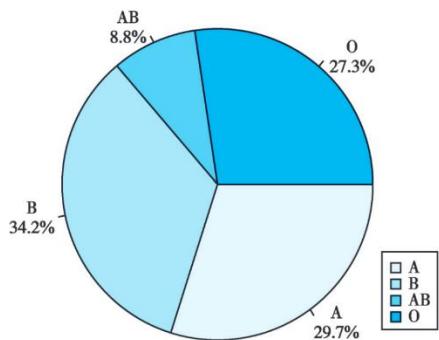


图 5-13 亚洲人群各种血型的构成比

(李康,贺佳主编, 2024, p. 53) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

绘制逻辑 (必须会)

- $360^\circ = 100\%$
- 某部分角度 = 构成比 $\times 3.6^\circ$

排列规则

- 从固定起点 (如 12 点方向)
- 顺时针
- 可按:
 - 数值大小
 - 或自然顺序

☛ 使用提醒 (非常重要)

- 圆图只适合:
 - 类别不多
 - 构成差异明显
- 类别一多、人眼就分不清了

2 百分条图 (Percent bar chart) ——看“多个整体的构成对比”

一句话本质

每一条都是 100%，重点是“不同整体的构成差异”

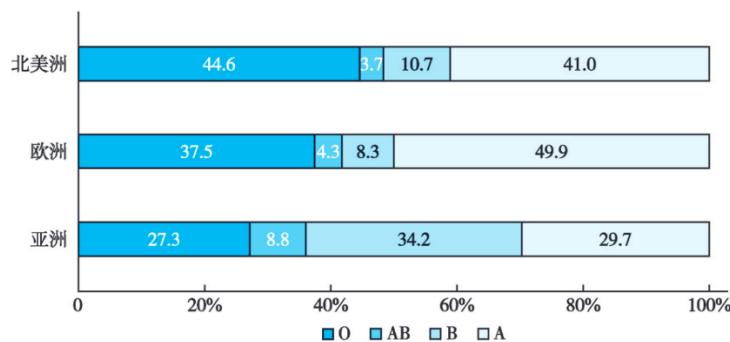


图 5-14 亚洲、欧洲和北美洲三个地区的人群血型分布

(李康, 贺佳主编, 2024, p. 53) - Please click Zotero - Refresh in Word/LibreOffice to update all fields.

相比圆图的优势

- 可并排比较多个群体

- 人眼对“长度”比“角度”更敏感
- 构成对比更清晰

👉 典型用途：

不同地区 / 不同人群的构成差异（如血型分布）

直条图 vs 构成图（必须分清）

问题类型	正确图形
谁更多 / 谁更少	直条图
总体中各部分占比	圆图 / 百分条图
不同群体构成对比	百分条图
绝对量比较	直条图
比例结构展示	构成图

👉 错误示范：

用圆图比较“哪个地区人数更多” —— ✗

用直条图展示“构成比” —— 容易误导

[BIBLIOGRAPHY] Please click Zotero - Refresh in Word/LibreOffice to update all fields