■ 表 8.5.3　χ^2 检验计算表

A_i	n_i	$\widehat{p}_i$	$n\widehat{p}_i$	$n_i^2/n\widehat{p}_i$
$A_1:-\infty<x\leqslant 2\ 300.5$	3	0.014 4	2.419	7.822 4
$A_2:2\ 300.5<x\leqslant 2\ 500.5$	5	0.034 3	5.762	
$A_3:2\ 500.5<x\leqslant 2\ 700.5$	13	0.080 9	13.591	12.435
$A_4:2\ 700.5<x\leqslant 2\ 900.5$	22	0.144 7	24.310	19.910
$A_5:2\ 900.5<x\leqslant 3\ 100.5$	28	0.197 2	33.130	23.665
$A_6:3\ 100.5<x\leqslant 3\ 300.5$	39	0.204 7	34.390	44.228
$A_7:3\ 300.5<x\leqslant 3\ 500.5$	28	0.161 6	27.149	28.878
$A_8:3\ 500.5<x\leqslant 3\ 700.5$	21	0.097 2	16.330	27.006
$A_9:3\ 700.5<x\leqslant 3\ 900.5$	7	0.044 5	7.470	10.922
$A_{10}:3\ 900.5<x<+\infty$	2	0.020 5	3.453	
合计				$\sum=174.866$

■ 思考题八

1. 如何理解小概率原理在假设检验中的应用?
2. 分别利用 $P-$值和显著水平 α 说明检验的统计显著性.
3. 有关参数的假设检验和分布的假设检验中, 如何根据样本资料合理地设置原假设和备择假设?
4. 将假设检验的原假设和备择假设互换, 如将 $H_0:\theta\leqslant\theta_0, H_1:\theta>\theta_0$ 换成 $H_0:\theta>\theta_0, H_1:\theta\leqslant\theta_0$, 检验结果有什么区别?
5. 对于一个假设检验问题, 设 $\alpha_1>\alpha_2$, 如果出现 "在显著水平 α_1 下拒绝原假设, 而在显著水平 α_2 下接受原假设", 这矛盾吗? 如何理解判断结果?
6. 假设检验问题中的两类错误是什么? 它们之间有什么关系? 它们和样本容量之间又有什么联系? 如何理解奈曼 - 皮尔逊原则?
7. 一家大型超市上个月接到许多消费者投诉某品牌土豆片 (标注 60 g/袋) 重量不符合. 店方猜想引起这些投诉的原因可能是运输过程导致产生很多的土豆片碎屑. 为了确定厂家的土豆片的重量是否不低于 60 g/袋, 店方决定对来自于一家最大的供应商的下一批袋装土豆片的平均重量 μ (单位: g) 进行检验, 假设陈述如下:

$$H_0:\mu\geqslant 60,\quad H_1:\mu<60.$$

　(1) 这一假设检验问题的第 I 类错误是什么?

　(2) 第 II 类错误是什么?

　(3) 你认为顾客会将哪类错误看得较为严重? 而供应商会将哪类错误看得较为严重?

8. 假设检验和区间估计有什么联系?

9. 说明皮尔逊拟合优度 χ^2 检验的基本思想.

▶ 习题八

1. 电视机显像管的质量标准是平均使用寿命为 15 000 h, 标准差为 1 500 h. 某电视机厂宣称其生产的显像管平均寿命大大高于规定的标准. 为了对此说法进行验证, 随机抽取了 100 件该厂产品为样本, 测得平均使用寿命为 15 525 h. 能否说该厂的显像管平均寿命显著地高于规定的标准? 取 $\alpha = 0.05$, 假设显像管的寿命 $X \sim N(\mu, 1500^2)$.

   (1) 给出检验的原假设和备择假设;

   (2) 求检验的拒绝域;

   (3) 求 $P-$值;

   (4) 根据样本判断是否可以认为该厂的显像管平均寿命显著高于规定的标准?

2. 某汽车厂商宣称他们生产的汽车平均每公升汽油可行驶 15 km 以上. 为验证该广告的真实性, 随机选取 10 辆车, 并且记录下每辆车每公升汽油行驶的千米数, 得到如下的观察值:

$$14.8 \quad 15.1 \quad 16.9 \quad 14.8 \quad 13.7 \quad 12.9 \quad 13.5 \quad 14.9 \quad 15.4 \quad 13.5$$

   假设数据来自正态分布, 请检验该广告的真实性 (取 $\alpha = 0.05$).

3. 设总体 $X \sim N(\mu, \sigma^2)$, 从总体中抽取容量为 16 的简单随机样本, 样本均值为 $\overline{X}$, 样本方差为 S^2.

   (1) 若 $\sigma^2 = 1$, 在显著水平为 0.05 下对于假设: $H_0 : \mu = 1, H_1 : \mu \neq 1$, 求出拒绝域; 并计算在 $\mu = 2$ 时犯第 II 类错误的概率;

   (2) 若 μ 未知, 在显著水平为 0.05 下对于假设: $H_0 : \sigma^2 = 1, H_1 : \sigma^2 > 1$, 求出拒绝域; 并计算在 $\sigma^2 = 4$ 时犯第 II 类错误的概率;

   (3) 若根据样本值得 $\overline{x} = 1.54, s^2 = 1.44$, 求 (1)、(2) 中的 $P-$值.

4. 火药生产厂家设计出一种新的火药生产方案, 要求使子弹发射的枪口速度达到 900 m/s. 假设枪口速度 X (m/s) 服从 $N(\mu, \sigma^2)$, 现做了 8 次试验, 其速度分别为

$$893 \quad 886 \quad 897 \quad 903 \quad 901 \quad 898 \quad 909 \quad 889$$

   (1) 试问这些数据是否足以说明其枪口速度的均值 μ与 900 m/s 存在显著差异 (取 $\alpha = 0.05$)?

   (2) 求 μ 的置信水平为 95% 的置信区间;

(3) 求 $P-$值.

5. 根据《中国居民营养与慢性病状况报告 (2015)》, 全国 18 岁及以上成年男性和女性的平均身高分别为 167.1 cm 和 155.8 cm, 平均体重分别为 66.2 kg 和 57.3 kg. 今从我国某东部地区随机抽选 400 名成年男子, 测得身高的平均值为 169.7 cm, 标准差为 4.2 cm. 设样本来自总体 $N(\mu,\sigma^2)$, 问该地区男子的身高是否明显高于全国平均水平 (取 $\alpha=0.05$).

6. 为调查某减肥药的成效, 随机选取 10 位使用者, 记录其服用减肥药前、后的体重 (单位: kg).

编号	1	2	3	4	5	6	7	8	9	10
服药前体重	66	70	56	58	49	75	63	56	48	75
服药后体重	68	65	54	59	45	70	60	50	47	68

试用假设检验方法分析该减肥药的效果 (取 $\alpha=0.05$).

7. 为了解某种犬类疫苗注射后是否会使得犬的体温升高, 随机选择 9 只狗, 记录它们注射疫苗前、后的体温 (单位: °C):

编号	1	2	3	4	5	6	7	8	9
注射前体温	37.5	37.7	38.1	37.9	38.3	38.5	38.1	37.6	38.4
注射后体温	37.7	38.0	38.2	37.9	38.2	38.8	38.0	37.5	38.8

设注射疫苗前、后体温差服从正态分布, 问是否可以认为注射疫苗后狗的体温有显著升高 ($\alpha=0.05$)?

8. 某经销代理商在和乳业公司的合约里要求生产的 225 mL 盒装牛奶中, 容量标准差不可超过 8 mL, 否则就予以退货. 现随机抽取 15 盒牛奶, 测得容量 (单位: mL) 分别为

$$\begin{array}{cccccccc} 230 & 223 & 228 & 229 & 220 & 215 & 217 & 231 \\ 220 & 223 & 230 & 224 & 226 & 228 & 227 & \end{array}$$

假设样本来自于正态总体, 在显著水平 $\alpha=0.05$ 下, 检验假设 $H_0:\sigma\geqslant 8, H_1:\sigma<8$, 并计算 $P-$值.

9. 某一灌装机灌装每瓶 550 mL 的饮料, 假设饮料的容量 X (单位: mL) 服从 $N(\mu,\sigma^2)$, 要求标准差不超过 5.5 mL. 为判断该灌装机工作是否正常, 从已灌装的饮料中随机取 9 瓶, 测得样本均值 $\overline{x}=553.5$, 样本标准差为 $s=6.3$. 在显著水平为 0.05 下,

(1) 检验假设 $H_0:\mu=550, H_1:\mu\neq 550$, 计算 $P-$值, 作出结论;

(2) $H_0:\sigma\leqslant 5.5, H_1:\sigma>5.5$, 计算 $P-$值, 作出结论.

10. 已知某种零件的长度 X (cm) 服从正态分布, 现从一批零件中随机抽取 16 只, 测得其长度如下:

$$\begin{array}{cccccccc} 15.1 & 14.9 & 14.8 & 14.6 & 15.2 & 14.8 & 14.9 & 14.6 \\ 14.8 & 15.1 & 15.3 & 14.7 & 15.0 & 15.2 & 15.1 & 14.7 \end{array}$$

(1) 若要求该种零件的标准长度应为 15 cm, 检验这批零件是否符合标准要求 (取 $\alpha = 0.05$);

(2) 若要求方差不超过 0.04, 问该批零件是否符合标准要求 (取 $\alpha = 0.05$).

11. 现对某论坛的日发帖量进行为期 2 周 (14 天) 的调查, 测得各日的日发帖量为 $x_1, \cdots, x_{14}$, 并计算得 $\sum\limits_{i=1}^{14} x_i = 7077, \sum\limits_{i=1}^{14}(x_i - \overline{x})^2 = 359.58$. 若假定该论坛的日发帖量 $X \sim N(\mu, \sigma^2)$.

(1) 通过计算 $P-$值说明是否有充分的理由认为 $\mu > 500$?

(2) 通过计算 $P-$值说明是否有充分的理由认为 $\sigma^2 \neq 30$ (取 $\alpha = 0.05$).

12. 下列数据为两个煤矿开采的每吨煤产生的热量记录 (单位: 4.186×10^3 J)

矿 A: 8 500　8 330　8 480　7 960　8 030

矿 B: 7 710　7 890　7 920　8 270　7 860

假设这些数据来自两个方差相等且相互独立的正态总体, 是否可以认为 A 矿的煤产生的热量要显著地大于 B 矿的煤 (取 $\alpha = 0.05$)?

13. 为比较甲、乙两位电脑打字员的出错情况, 随机抽查甲输入的文件 8 页, 各页出错字数为

$$5 \quad 3 \quad 2 \quad 0 \quad 1 \quad 2 \quad 2 \quad 4$$

抽查乙输入的文件 9 页, 各页出错字数为

$$5 \quad 1 \quad 3 \quad 2 \quad 4 \quad 6 \quad 4 \quad 2 \quad 5$$

假设甲、乙两人页出错字数都服从正态分布, 试检验

(1) 甲、乙两人页出错数的方差是否相等 (取 $\alpha = 0.05$)?

(2) 甲页均出错数是否显著少于乙 (取 $\alpha = 0.05$)?

14. 为了研究男性长跑运动员的心率是否低于一般健康男性心率, 现从省长跑队随机抽取了 10 名运动员, 并从某高校随机抽取 25 名学生. 测得运动员心率的平均值为 60 次/分, 标准差为 6 次/分, 大学生心率的平均值为 73 次/分, 标准差为 13 次/分. 假设心率次数服从正态分布, 根据上面的资料检验

(1) 两个群体心率的方差是否相等 (取 $\alpha = 0.05$)?

(2) 若假设两个群体心率的方差不相等, 是否有理由认为男性长跑运动员每分钟的心率显著低于一般年轻男性 (取 $\alpha = 0.05$)?

15. 为研究某种新药对抗凝血酶活力的影响, 随机安排新药组病人 12 例, 对照组病人 10 例, 分别测定其抗凝血酶活力 (单位: mm^3), 假设数据来自正态总体, 结果如下:

新药组: 136　127　128　128　133　138　142　116　110　108　115　140

对照组: 163　165　177　170　175　152　157　159　160　164

试在显著水平 0.05 下检验

(1) 两个总体的方差是否相等?

(2) 新药组和对照组病人的平均抗凝血酶活力有无显著差异?

16. 一盒中有 10 个球, 其中红球有 a 个 (未知), 其余是白球, 采用放回抽样取 3 个球作为一次试验, 这样的试验总共进行 200 次, 发现有 40 次没有取到红球, 有 85 次取到 1 个红球, 有 63 次取到 2 个红球, 有 12 次取到 3 个红球, 请在显著水平 0.05 下, 检验假设 $H_0: a = 3$.

17. 某个八面体各面分别标有数字 1, 2, 3, 4, 5, 6, 7, 8, 为检验它各面是否匀称, 即各面出现的概率是否均相等, 作 600 次投掷试验, 各数字朝上的次数如下:

数字:	1	2	3	4	5	6	7	8
频数:	72	83	78	90	70	71	64	72

在显著水平 0.05 下, 检验假设 H_0: 该八面体是匀称的.

18. 对某公交车站观察从 12 点到 15 点这 3 个小时前来等车的乘客情况, 将 2 分钟作为一个单位时间, 记录 90 个单位时间等车的情况, 数据如下:

乘客数:	0	1	2	3	4	5	7	8
频数:	5	12	18	21	16	13	3	2

问这些数据是否来自泊松分布的总体 (取 $\alpha = 0.05$)?

19. 某 ATM 机等候一个顾客来到的时间为 X (单位: min), 现观察了 100 次, 获得如下数据:

等候时间 X:	$0 \leqslant X \leqslant 5$	$5 < X \leqslant 10$	$10 < X \leqslant 20$	$20 < X \leqslant 30$	$X > 30$
频数:	30	28	20	15	7

在显著水平 0.05 下, 检验假设 H_0: 等候时间 X 服从均值为 10 的指数分布.

20. 对某地区成年男子身高 X (单位: cm) 进行观察, 随机抽取 200 名男子, 得到样本均值和样本标准差分别为 $\overline{x} = 169.9, s = 9.6$, 其他资料如下:

身高 x:	$x \leqslant 163$	$163 < x \leqslant 167$	$167 < x \leqslant 171$	$171 < x \leqslant 175$	$175 < x \leqslant 179$	$179 < x \leqslant 183$	$x > 183$
频数:	41	34	40	33	27	16	9

在显著水平 0.05 下, 检验假设 H_0: 该地区成年男子身高服从正态分布.