12.23 第九章课后习题：1、3、5、7、9、10、12

1. 为检验吸烟与患慢性气管炎有无关系，随机调查了 339 人，其中 205 名吸烟者中有 43 人患慢性气管炎，在 134 名不吸烟者中有 13 人患慢性气管炎. 问在显著水平 0.05 下数据是否支持“吸烟者患慢性气管炎的比例较高”这个结论？

Chap 1.

1. 解: H_0：吸烟与患慢性气管炎独立.

	不吸烟	吸烟	
无气管炎	121	162	283
患气管炎	13	43	56
	134	205	339

$$Z=\frac{n(n_{11}n_{22}-n_{12}n_{21})^2}{n_{1\cdot}n_{2\cdot}n_{\cdot 1}n_{\cdot 2}}$$

$$Z=\frac{339(121\times 43-162\times 13)^2}{283\times 56\times 134\times 205}$$

$$\approx 7.4688.$$

$\chi^2_{(a-1)(b-1)}=\chi^2_1$　$\chi^2_1(0.05)=3.841$　$\because Z>\chi^2_1(0.05)$ 拒绝 H_0.

$\therefore$ 吸烟与患慢性气管炎不独立. 吸烟者患慢性气管炎比例较高.

3.一农场半年前在一鱼塘中按比例 20：15:40:25 投放了四种鱼：鲑鱼、鲈鱼、竹夹鱼和鲇鱼的鱼苗，现在在鱼塘里获得一个样本如下：

序号	1	2	3	4
种类	鲑鱼	鲈鱼	竹夹鱼	鲇鱼
数量/条	132	100	200	168

试取 $\alpha=0.05$, 检验各类鱼数量的比例较半年前是否有显著的改变 ：

3. 解: $F:\begin{pmatrix}1 & 2 & 3 & 4\\ 0.2 & 0.15 & 0.4 & 0.25\end{pmatrix}$　$n=\sum n_i=600.$

H_0: 各类鱼数量比例较半年前未改变 / X(属于哪一类鱼) $\sim F$.

$$Z=\sum_{i=1}^{4}\frac{(np_i-n_i)^2}{np_i}=\frac{(120-132)^2}{120}+\frac{(90-100)^2}{90}+\frac{(240-200)^2}{240}+\frac{(150-168)^2}{150}\approx 11.1378.$$

$\chi^2_{4-1}(0.05)=7.815$　$\because Z>\chi^2_3(0.05)$ 拒绝 H_0.

$\therefore$ 认为各类鱼比例较半年前有变化.

5. 某媒体用抽样调查来确定人们如何利用他们的空闲时间. 男性和女性都选择看电视为最普遍的活动，调查结果如下表所示：

性别		样本容量	选择看电视人数
男性		800	248
女性		600	156

(1)陈述一个假设，该假设可用于检验选择看电视作为最普遍活动男性比率和女性比率之间的差异.

(2)选择看电视作为最普遍活动的男性样本比率是多少？相应的女性样本比率是多少？

(3)进行假设检验并计算 p 值. 在显著性水平 $\alpha = 0.05$ 下，你的结论如何？

(4)总体比率之差的 95% 置信区间估计是什么？

5.解：(1) H_0：男性与女性选择看电视为最普遍活动的比率相同.

(2) 男性：$\frac{248}{800} = 31\%$　　女性：$\frac{156}{600} = 26\%$

(3)

	选看电视.	未选~	
男	248	552	800
女	156	444	600
	404	996	1400

$$\therefore Z = \frac{1400(248\times444-552\times156)^2}{800\times600\times404\times996} \approx 4.1751 > \chi_1^2(0.05) = 3.841.$$

p值 $= P(\chi_1^2 > 4.1751) \approx 0.041 < 0.05$　(用软件计算、考试肯定不会考计算p值).

$\therefore$ 拒绝 H_0. 男性和女性选择看电视为最普遍活动的比率有差异.

(4) 假设调查的所有人之间相互独立. 男性选择看电视: $X \sim B(1,p_1)$ 样本另 $n_{1\cdot}=800$

女性 . . . . : $Y \sim B(1,p_2)$. . . . $n_{2\cdot}=600$

题目中比率之差应为概率之差 p_1-p_2.

利用大样本方法.

$$\frac{\bar{X}-\bar{Y}-(p_1-p_2)}{\sqrt{Var(\bar{X}-\bar{Y})}} \xrightarrow{d} N(0,1). \quad as\ n\to\infty.$$

$$Var(\bar{X}-\bar{Y}) = Var\bar{X}+Var\bar{Y} = \frac{VarX}{n_{1\cdot}}+\frac{VarY}{n_{2\cdot}} = \frac{p_1(1-p_1)}{n_{1\cdot}}+\frac{p_2(1-p_2)}{n_{2\cdot}}$$

p_1. p_2 的 MLE 分别为 $\hat{p}_1=\frac{n_{11}}{n_{1\cdot}}=\frac{248}{800}=0.31$. $\hat{p}_2=\frac{n_{21}}{n_{2\cdot}}=\frac{156}{600}=0.26$.

$$\frac{\frac{\hat{p}_1(1-\hat{p}_1)}{n_{1\cdot}}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_{2\cdot}}}{\sqrt{Var(\bar{X}-\bar{Y})}} \xrightarrow{P} 1 \qquad \therefore \frac{\bar{X}-\bar{Y}-(p_1-p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_{1\cdot}}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_{2\cdot}}}} \xrightarrow{d} N(0,1).$$

$$T=\frac{0.31-0.26-(p_1-p_2)}{\sqrt{\frac{0.31(1-0.31)}{800}+\frac{0.26(1-0.26)}{600}}} \xrightarrow{d} N(0,1)$$

令 $-u_{\frac{\alpha}{2}} \le T \le u_{\frac{\alpha}{2}}$ $\quad u_{0.025}=1.96$.

$\therefore$ 解得 $0.00247 \le p_1-p_2 \le 0.0975$. $\therefore$ 95% 置信区间为 $[0.00247, 0.0975]$

7. 摩尔根的果蝇实验用来检验孟德尔第二定律 (自由组合定律) 是否成立.在该定律成立的条件下，果蝇眼睛的颜色 (红-A,紫-a) 和翅膀的长度 (长-B,短-b) 应该是独立遗传的，即有4种组合 (AB, Ab, aB, ab) 应该是等可能的 (概率各为0.25).摩尔根观察到的4种组合的计数分别为 1339，151，154，1195 (总数 n=2839). 试检验孟德尔第二定律是否成立.

7. 解: $H_0: P(AB)=P(Ab)=P(aB)=P(ab)=\frac{1}{4}$. (孟德尔第二定律成立).

$np_i = 2839\times\frac{1}{4}$.

$$Z=\sum_{i=1}^{4}\frac{(np_i-n_i)^2}{np_i} = 1764.682 \gg \chi^2_3(0.05)$$

$\therefore$ 拒绝 H_0. 孟德尔第二定律不成立.

10. 有甲、乙、丙三个工厂生产同一种产品。产品分一，二, 三个等级 (分别代表高、中、低). 为考察各工厂产品质量是否一致，从这三个工厂中分别随机抽出产品若干件，每件鉴定其质量等级，结果如下：

等级	甲	乙	丙
一	58	38	30
二	40	44	35
三	11	18	26

10. 解. 齐一性检验.

H_0: 各个工厂产品质量一致.

$$Z=\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{\left(n_{ij}-\frac{n_{i\cdot}n_{\cdot j}}{n}\right)^2}{\frac{n_{i\cdot}n_{\cdot j}}{n}}\approx 15.407>\chi^2_{2\times2}(0.05)=9.488$$

∴拒绝H_0. 各工厂产品质量不一致.

	甲	乙	丙	
一	58	38	30	126
二	40	44	35	119
三	11	18	26	55
	109	100	91	300

从各工厂产品不同等级的比例能看出甲厂较优，丙厂较劣.

一等品：甲：0.532 > 乙：0.38 > 丙：0.33

二等品 甲：0.367 < 乙：0.44 > 丙：0.385

三等品 甲：0.1 < 乙：0.18 < 丙：0.286

9. 为了研究蜗牛的种类是否与其生活的珊瑚礁种类有关，选取了 3 种珊瑚礁作为检验样本，记为 I, II, III , 记录下 A 和 B 两种蜗牛分别在 3 种珊瑚礁中生存的数目，得到如下数据.试问 A 和 B 两种蜗牛的分布是否在 3 种珊瑚礁中都是一样的($\alpha = 0.05$)?

	I	II	II	合计
A	6	8	14	28
B	7	21	5	33
合计	13	29	19	61

解：方法与第 10 题一致。

$$Z = \sum_{i=1}^{2}\sum_{j=1}^{3}\frac{(n_{ij} - \frac{n_i - n_{ij}}{n_{ij}})^2}{\frac{n_i \cdot n_{ij}}{n}} = 4.397 < \chi_2^2(0.05) = 5.99$$

故可以认为 A 和 B 两种蜗牛的分布是否在 3 种珊瑚礁中都是一样的.

12. 为了解男性和女性对三种类型的啤酒：淡啤酒、普通啤酒和黑啤酒的偏好有没有差异分别调查了 180 位男性和 120 位女性的喜好，得如下数据： 请问男性和女性对这三种类型的啤酒的偏好有显著差异吗？($\alpha = 0.05$)

	淡啤酒	普通啤酒	黑啤酒
男性	49	31	100
女性	51	20	49

12. 解： H_0: 男性和女性对三类啤酒偏好无差异.

	淡	普通	黑.	
男	49 (60)	31	100	180
女	51	20	49	120
	100	51	149	300

$$Z = \sum_{i=1}^{2}\sum_{j=1}^{3} \frac{\left(n_{ij} - \frac{n_{i\cdot}n_{\cdot j}}{n}\right)^2}{\frac{n_{i\cdot}n_{\cdot j}}{n}} \approx 8.1968 > \chi^2_{2\cdot 1}(0.05) = 5.991$$

∴ 拒绝 H_0. 偏好有显著差异.