第 2 节 独立性检验（★★☆）

强化训练

1．（2023·上海模拟·★★）某地政府调查育龄妇女生育意愿与家庭年收入高低的关系时，随机调查了当地 3000 名育龄妇女，用独立性检验的方法处理数据，并计算得 $\chi^2 = 7.326$，则根据这一数据以及临界值表，判断育龄妇女生育意愿与家庭年收入高低有关系的可信度（　　）

（A）低于 1%　　（B）低于 0.5%　　（C）高于 99%　　（D）高于 99.5%

参考数据：$P(\chi^2 \geq 10.828) \approx 0.001$，$P(\chi^2 \geq 7.879) \approx 0.005$，$P(\chi^2 \geq 6.635) \approx 0.01$，$P(\chi^2 \geq 3.841) \approx 0.05$，$P(\chi^2 \geq 2.706) \approx 0.1$.

答案：C

解析：因为 $6.635 < \chi^2 = 7.326 < 7.879$，且 $P(\chi^2 \geq 6.635) \approx 0.01$，$P(\chi^2 \geq 7.879) \approx 0.005$，所以判断育龄妇女生育意愿与家庭年收入高低有关系，这种判断出错的概率小于 0.01，即可信度高于 99%，但不足 99.5%.

2．（2023·云南统考·★★）党的二十大胜利召开后，某校为调查性别因素对党史知识的了解情况是否有影响，随机抽查了男女教职工各 100 名，得到如下数据：

	不了解	了解
女职工	30	70
男职工	20	80

根据小概率值 $\alpha = 0.005$ 的独立性检验，能否认为对党史知识的了解情况与性别有关？

参考公式：$\chi^2 = \dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中 $n = a+b+c+d$.

附表：

α	0.010	0.005	0.001
x_α	6.635	7.879	10.828

解：（第一步，写出零假设）零假设为 H_0：对党史知识的了解情况与性别无关，

（第二步，由列联表计算 χ^2 的值，并与表中对应的 α 临界值比较）

$$\chi^2 = \frac{200\times(30\times80-20\times70)^2}{50\times150\times100\times100} \approx 2.667 < 7.879,$$

（第三步，结合比较的结果作答）根据小概率值 $\alpha = 0.005$ 的独立性检验，没有充分的证据推断 H_0 不成立，所以可以认为对党史知识的了解情况与性别无关.

3．（2023·湖北模拟·★★★）某数学兴趣小组为研究本校学生数学成绩与语文成绩的关系，采取有放回的简单随机抽样，从学校抽取样本量为 200 的样本，将所得数学成绩与语文成绩的样本观测数据整理如下：

	语文成绩		合计
	优秀	不优秀	

数学成绩	优秀	50	30	80
	不优秀	40	80	120
合计		90	110	200

（1）根据小概率值$\alpha = 0.01$的独立性检验，能否认为数学成绩与语文成绩有关联？

（2）在人工智能中常用$L(B\,|\,A) = \dfrac{P(B\,|\,A)}{P(\bar{B}\,|\,A)}$表示在事件$A$发生的条件下事件$B$发生的优势，在统计中称为似然比. 现从该校学生中任选一人，A表示“选到的学生语文成绩不优秀”，B表示“选到的学生数学成绩不优秀”，请利用样本数据，估计$L(B\,|\,A)$的值.

参考公式：$\chi^2 = \dfrac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$，其中$n = a+b+c+d$.

附表：

α	0.01	0.005	0.001
x_α	6.635	7.879	10.828

解：（1）（第一步，写出零假设）零假设为H_0：数学成绩与语文成绩无关联，

（第二步，由列联表计算χ^2的值，并与表中对应的α临界值比较）

$$\chi^2 = \frac{200\times(50\times80-30\times40)^2}{90\times110\times80\times120} \approx 16.498 > 6.635,$$

（第三步，结合比较的结果作答）根据小概率值$\alpha = 0.01$的独立性检验，我们推断H_0不成立，所以能认为数学成绩与语文成绩有关联.

（2）由题意，$L(B\,|\,A) = \dfrac{P(B\,|\,A)}{P(\bar{B}\,|\,A)} = \dfrac{\dfrac{P(AB)}{P(A)}}{\dfrac{P(A\bar{B})}{P(A)}} = \dfrac{P(AB)}{P(A\bar{B})}$ ①，（要求此式的值，可用样本数据来估算$P(AB)$和$P(A\bar{B})$）

由列联表可知$P(AB)$的估计值为$\dfrac{80}{200}$，$P(A\bar{B})$的估计值为$\dfrac{30}{200}$，代入①得$L(B\,|\,A)$的估计值为$\dfrac{8}{3}$.