概率论第八章 答案

1. 设 X 是连续随机变量，x 是对 X 的（一次）观测值。关于 X 的密度函数 $f(x)$ 由如下两个假设

$$H_0: f(x)=\begin{cases}\dfrac{1}{2}, & 0\leqslant x\leqslant 2\\ 0, & \text{其他}\end{cases}; H_1: f(x)=\begin{cases}\dfrac{x}{2}, & 0\leqslant x\leqslant 2\\ 0, & \text{其他}\end{cases}$$

检验的判断规则是：若 $x\geqslant\frac{2}{3}$ 则拒绝 H_0，试求犯两类错误的概率。

解 根据所给条件，犯第一类错误的概率为

$$\alpha=P\left(x\geqslant\frac{2}{3}|H_0\right)=\int_{\frac{2}{3}}^{2}\frac{1}{2}\mathrm{d}x=\frac{2}{3}.$$

犯第二类错误的概率为

$$\beta=P\left(x<\frac{2}{3}|H_1\right)=\int_{0}^{\frac{2}{3}}\frac{x}{2}\mathrm{d}x=\frac{1}{9}.$$

2. 设 $X_1,X_2,\cdots,X_{36}$ 是来自正态总体 $N(\mu,0.04)$ 的一个简单随机样本，其中 μ 未知，记 $\bar{X}=\dfrac{1}{36}\displaystyle\sum_{i=1}^{36}X_i$，现对检验问题 $H_0:\mu=0.5$，$H_1:\mu=\mu_1>0.5$，并取拒绝域 $D=\{(X_1,X_2,\cdots,X_{36}):\bar{X}>C\}$，显著性水平 $\alpha=0.05$。

(1) 求常数 C;

(2) 若 $\alpha=0.05$，$\mu_1=0.65$ 时，犯第二类错误的概率是多少?

解 (1) 若假设 H_0 成立，即 $\mu=0.5$，那么总体 $X\sim N(0.5,0.04)$，$\bar{X}\sim N(0.5,\dfrac{1}{900})$.

根据题意知

$$\begin{aligned}\alpha&=P\{\text{拒绝}H_0|H_0\text{为真}\}\\&=P\{\bar{X}>C\}\\&=1-P\{\bar{X}\leqslant C\}\\&=1-P\left\{\frac{\bar{X}-0.5}{1/30}\leqslant\frac{C-0.5}{1/30}\right\}\\&=1-\Phi(30C-15)\\&=0.05\end{aligned}$$

查表得 $30C-15=1.645$，即 $C=0.5548$.

(2) 若假设 H_1 成立，即 $\mu=\mu_1=0.65$，那么总体 $X\sim N(0.65,0.04)$，$\bar{X}\sim N(0.65,\dfrac{1}{900})$.

根据题意知

$$
\begin{aligned}
\beta &= P\{\text{接受}H_0|H_0\text{不真}\} \\
&= P\{\bar{X} < C\} \\
&= P\left\{\frac{\bar{X}-0.65}{1/30} < \frac{C-0.65}{1/30}\right\} \\
&= \Phi[30\times(0.5548-0.65)] \\
&= 1-\Phi(2.855) \\
&= 0.0021.
\end{aligned}
$$

3. 已知维尼纶纤度在正常情况下服从正态分布，且标准差为 0.048。从某天产品中抽取 5 根纤维，测得其纤度分别为

$$1.32 \quad 1.55 \quad 1.36 \quad 1.40 \quad 1.44$$

在显著性水平为 0.05 下，问这一天纤度的总体标准差是否正常?

解 这是关于正太总体方差的双侧假设检验问题:

$$H_0:\sigma^2=0.048^2 \quad H_1:\sigma^2\neq 0.048^2$$

$n=5$，$\alpha=0.05$，查表得 $\chi^2_{0.975}(4)=0.4844$，$\chi^2_{0.025}(4)=11.1433$，于是检验的拒绝域为 $W=\{\chi^2\leqslant 0.4844\text{或}\chi^2\geqslant 11.1433\}$.

根据样本数据，计算得到

$$\chi^2=\frac{(n-1)s^2}{\sigma_0^2}=13.5069>11.1433=\chi^2_{0.025}(4).$$

因此在 $\alpha=0.05$ 时，拒绝原假设，可以认为这一天纤度的总体标准差不正常。

4. 为检验一颗骰子是否均匀，将它投掷 60 次，观测到出现点数 1,2,3,4,5,6 的次数分别为 7,6,12,14,5,16，问这颗骰子是否均匀?（取显著水平 $\alpha=0.05$）

解 用 X 表示出现的点数，则 X 是离散型随机变量，检验假设

$$H_0:P\{X=i\}=\frac{1}{6} \quad (i=1,2,3,4,5,6)$$

这里 $r=6,t=0,n=60,p_i=\frac{1}{6}(i=1,2,3,4,5,6)$。计算结果列入表格如下:

i	f_i	p_i	np_i	f_i-np_i	$\frac{(f_i-np_i)^2}{np_i}$
1	7	$\frac{1}{6}$	10	-3	0.9
2	6	$\frac{1}{6}$	10	-4	1.6
3	12	$\frac{1}{6}$	10	2	0.4
4	14	$\frac{1}{6}$	10	4	1.6
5	5	$\frac{1}{6}$	10	-5	2.5
6	16	$\frac{1}{6}$	10	6	3.6
χ^2					10.6

由附表差得 $\chi^2_\alpha(r-1)=\chi^2_{0.05}(5)=11.071>10.6$，因此，接受 H_0，认为这颗骰子是均匀的。

5. 下表分别给出两位文学家马克 · 吐温 (Mark Twain) 的 8 篇小品文以及斯诺特格拉斯 (Snodgrass) 的 10 篇小品文中由 3 个字母组成的单字的比例:

马克 · 吐温	0.225	0.262	0.217	0.240	0.230	0.229	0.235	0.217	
斯诺特格拉斯	0.209	0.205	0.196	0.210	0.202	0.207	0.224	0.223	0.220	0.201

设两组数据分别来自正态总体, 且两总体方差相等, 但参数均未知. 两样本相互独立. 问两位作家所写的小品文中包含由 3 个字母组成的单字的比例是否有显著的差异 (取 $\alpha=0.05$)?

解 按题意总体 $X \sim N\left(\mu_1, \sigma_1^2\right), Y \sim N\left(\mu_2, \sigma_2^2\right)$, 两总体的方差相等, 均等于 σ^2, σ^2 未知，两样本相互独立, 本题需在显著性水平 $\alpha=0.05$ 下检验假设

$$H_0: \mu_1=\mu_2, \quad H_1: \mu_1 \neq \mu_2$$

采用 t 检验, 取检验统计量为 $t=\frac{\bar{X}-\bar{Y}-\delta}{S_w \sqrt{1 / n_1+1 / n_2}}$, 今 $n_1=8, n_2=10, \bar{x}=0.2319, \bar{y}=0.2097, s_1^2=0.0146^2, s_2^2=0.0097^2, s_w^2=\frac{(8-1) s_1^2+(10-1) s_2^2}{8+10-2}=0.012^2, \delta=0, t_{0.025}(16)=2.1199$. 拒绝域为

$$|t|=\left|\frac{\bar{x}-\bar{y}}{s_w \sqrt{1 / n_1+1 / n_2}}\right| \geqslant t_{a / 2}\left(n_1+n_2-2\right)=2.1199$$

因观察值

$$|t|=\left|\frac{0.2319-0.2097}{0.012 \sqrt{1 / 8+1 / 10}}\right|=3.900>2.1199$$

落在拒绝域之内, 故拒绝 H_0, 认为两位作家所写的小品文中包含由 3 个字母组成的单字的比例有显著的差异.

6. 冶炼某种金属有两种方法，为检验用这两种方法生产的产品中所含杂质的波动性是否有明显的差异，各取一个样本，得数据（含杂质的百分数）如下：

甲 26.9 22.8 25.7 23.0 22.3 24.2 26.1

26.4 27.2 30.2 24.5 29.5 25.1

乙 22.6 22.5 20.6 23.5 24.3 21.9 20.6

23.2 23.4

由经验知道，产品中的杂质含量服从正态分布（取 $\alpha=0.05$）

解 设甲、乙两种冶炼方法所生产的产品中杂质含量分别为 X，Y，则 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$。检验杂质含量的波动性的大小，也就是比较总体方差的大小。因此提出检验假设：

$$\mathrm{H}_0: \sigma_1^2=\sigma_2^2 ; \mathrm{H}_1: \sigma_1^2 \neq \sigma_2^2$$

当 H_0 为真时

$$F=\frac{S_1^2}{S_2^2} \sim F\left(n_1-1, n_2-1\right)$$

由已知条件算得

$$n_1=13, n_2=9, \bar{x}=25.26, \bar{y}=22.51, s_1^2=5.862, s_2^2=1.641$$

$$F_0=\frac{s_1^2}{s_2^2}=3.572<F_{\frac{s}{2}}\left(n_1-1, n_2-1\right)=F_{0.025}(12,8)=4.20$$

所以接受 H_0，认为甲乙两种方法所生产的产品杂质含量波动性无显著差异。

若考虑作单侧检验，提出检验假设:

$$\mathrm{H}_0:\sigma_1^2=\sigma_2^2;\mathrm{H}_1:\sigma_1^2>\sigma_2^2$$

当 H_0 为真时，算得

$$F_0=\frac{s_1^2}{s_2^2}=3.572>F_{0.05}(12,8)=3.28$$

则应该拒绝 H_0，接受 H_1，即认为甲种方法生产的杂质含量的波动性比较大。

7. 随机抽取 200 只某种电子元件进行寿命试验，测得元件的寿命（单位：h）的频数分布表如下表:

元件寿命	≤ 200	$(200,300]$	$(300,400]$	$(400,500]$	>500
频数	94	25	22	17	42

根据计算，平均寿命为 325h，试检验元件的寿命是否服从指数分布? (取显著水平 $\alpha=0.10$)

解 由题可得极大似然估计位置参数 λ，有 $\hat{\lambda}=\frac{1}{\bar{X}}$，平均寿命 $\bar{X}=325$，因此原假设可以改写为 $H_0:X\sim E\left(\frac{1}{325}\right)$，则每个区间的理论概率值为 $\hat{P}_i=P(x_i<X\leq x_{i+1})=F(x_{i+1})-F(x_i)$，进而可计算出如下结果:

元件寿命区间	频数	概率	理论频数	G
<200	94	0.4596	91.9	0.048
$(200,300]$	25	0.1431	28.6	0.453
$(300,400]$	22	0.1052	21	0.048
$(400,500]$	17	0.0774	15.5	0.145
>500	42	0.2147	42.9	0.019
Σ	200	1		0.713

其中概率为 $\hat{P}_i$，理论频数为 $n\times\hat{P}_i$，$G=\frac{(n_i-n\times\hat{P}_i)^2}{n\hat{P}_i}$。由此可得 $\chi^2=\sum_{i=1}^{5}\frac{(n_i-n\times\hat{P}_i)^2}{n\hat{P}_i}=0.713$。由于估计了 1 个参数，所以 $\chi^2\sim\chi^2(3)$，对显著性水平 $\alpha=0.1$，查表得 $\chi^2_{0.1}(3)=6.252$，故此可得 H_0 的拒绝域 $W=(\chi^2>6.252)$，显然检验统计量不在拒绝域内，所以接受原假设，即认为该电子元件使用寿命的分布是指数分布 $H_0:X\sim E\left(\frac{1}{325}\right)$。

8. 随机地选了 8 个人，分别测量了他们在早晨起床时和晚上就寝时的身高（cm），得到以下的数据:

序号	1	2	3	4	5	6	7	8
早上(x_i)	172	168	180	181	160	163	165	177
晚上(y_i)	172	167	177	179	159	161	166	175

设各对数据的差 $D_i=x_i-y_i$ $(i=1,2,\ldots,8)$ 是来自正态总体 $N(\mu_D,\sigma_D^2)$ 的样本，μ_D,σ_D^2 均未知。问是否可以认为早晨的身高比晚上的身高更高（取 $\alpha=0.05$)?

解 题中的数据属成对数据，且可以认为成对数据之差来自正态总体 $N(\mu_D,\sigma_D^2)$，本题要求在显著性水平 $\alpha=0.05$ 下检验假设:

$$H_0:\mu_D\leq 0,\quad H_1:\mu_D>0$$

采用 t 检验，取检验统计量：

$$t = \frac{\bar{D} - 0}{S_D/\sqrt{n}}, \quad n = 8$$

各对数据之差 $d_i = x_i - y_i$ $(i = 1, 2, \ldots, 8)$ 依次为：

$$0,\ 1,\ 3,\ 2,\ 1,\ 2,\ -1,\ 2$$

由此得：

$$\bar{d} = \frac{1}{8}\sum_{i=1}^{8} d_i = 1.25, \quad s_D = 1.2817$$

查 t 分布表得：

$$t_\alpha(7) = t_{0.05}(7) = 1.8946$$

拒绝域为：

$$t = \frac{\bar{d} - 0}{s_D/\sqrt{n}} > t_\alpha(n-1) = 1.8946$$

计算观察值：

$$t = \frac{1.25 - 0}{1.2817/\sqrt{8}} = 2.758 > 1.8946$$

落在拒绝域之内，故在显著性水平 $\alpha = 0.05$ 下拒绝 H_0，认为早晨的身高比晚上的身高要高。

9. 某食品厂用自动装罐机装罐头，规定其标准重量为 250g，标准差不超过 3g 时判定该机器工作正常，每天定时检验机器的工作情况。现抽取 16 罐，测得平均重量 $\bar{x} = 252$g，样本标准差 $s = 4$g。假定罐头重量服从正态分布，试问该机器目前的工作是否正常？(取显著水平 $\alpha = 0.05$)

解 依题意，罐头重量 $X \sim N(\mu, \sigma^2)$，$n = 16$，$\bar{x} = 252$，$s = 4$，$\alpha = 0.05$。先检验 μ 是否等于 250g，再检验 σ^2 是否超过了 3^2。

(1) 先检验 μ 是否等于 250g。按题意总体 $X \sim N(\mu, \sigma^2)$，μ，σ^2 均未知，要求在水平 $\alpha = 0.05$ 下检验假设 $H_0 : \mu = 250$，$H_1 : \mu \neq 250$。

由于因 σ^2 未知，故采用 t 检验，取检验统计量为 $t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$，又 $t_{\frac{\alpha}{2}}(n-1) = t_{0.025}(15) = 2.1315$，拒绝域为：

$$|t| = \left|\frac{\bar{x} - \mu}{s/\sqrt{n}}\right| \geq t_{0.025}(15) = 2.1315$$

由于观察值 $|t| = \left|\frac{252-250}{\frac{4}{\sqrt{16}}}\right| = 2 < 2.1315$，没落在拒绝域内，所以接受 H_0。

(2) 检验 σ^2 是否超过了 3^2。检验假设 $H_0' : \sigma^2 \leq \sigma_0^2 = 9$，$H_1' : \sigma^2 > \sigma_0^2 = 9$。

因为 μ 未知，所以选取统计量 $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$，且 $\chi^2_\alpha(15) = \chi^2_{0.05}(15) = 24.996$，拒绝域为：

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} > \chi^2_{0.05}(15) = 24.996$$

由于 $\chi^2 = \frac{15}{9} \times 4^2 = 26.667 > \chi^2_{0.05}(15) = 24.996$，所以拒绝 H_0'。

综合 (1) 和 (2)，在显著性水平 $\alpha = 0.05$ 时，可以认为机器工作不正常。

10. 一种混杂的小麦品种，株高的标准差为 $\sigma_0 = 14$ cm，经提纯后随机抽取 10 株，它们的株高（以 cm 计）为：

$$90 \quad 105 \quad 101 \quad 95 \quad 101 \quad 100 \quad 101 \quad 105 \quad 93 \quad 97$$

考察提纯后群体是否比原群体整齐？取显著性水平 $\alpha = 0.01$，并设小麦株高服从 $N(\mu, \sigma^2)$。

解 需检验假设（$\alpha = 0.01$）：

$$H_0 : \sigma \geq \sigma_0, \quad H_1 : \sigma < \sigma_0 \quad (\sigma_0 = 14)$$

采用 χ^2 检验法。拒绝域为：

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{1-\alpha}(n-1)$$

现在，$n = 10$，$\chi^2_{1-0.01}(9) = 2.088$，$s^2 = 24.233$，
代入得：

$$\frac{(n-1)s^2}{\sigma_0^2} = \frac{218.1}{14^2} = 1.11 < 2.088$$

故拒绝 H_0，认为提纯后的群体比原群体整齐。