

第五章 数理统计初步习题参考答案与提示

1. 在总体 $X \sim N(52, 6.3^2)$ 中随机抽取一长度为 36 的样本, 求样本均值 \bar{X} 落 50.8 到 53.8 之间的概率。

答案与提示: 由于 $\bar{X} \sim N(\mu, \sigma^2/n)$, 所以

$$P\{50.8 < \bar{X} < 53.8\} = 0.8293.$$

2. 在总体 $X \sim N(8, 20^2)$ 中随机抽取一长度为 100 的样本, 问样本均值与总体均值的差的绝对值大 3 的概率是多少?

答案与提示: 由于 $\bar{X} \sim N(\mu, \sigma^2/n)$, 所以

$$P\{|\bar{X} - 8| > 3\} = 0.1336$$

3. 设 X_1, X_2, \dots, X_n 为来自总体 $X \sim P(\lambda)$ 的一个样本, \bar{X} 、 S^2 分别为样本均值和样本方差。求 $D\bar{X}$ 及 ES^2 。

答案与提示: 此题旨在考察样本均值的期望、方差以及样本方差的期望与总体期望、总体方差的关系, 显然应由定理 5-1 来解决这一问题。

$$D\bar{X} = \frac{DX}{n} = \frac{\lambda}{n}, \quad ES^2 = \lambda.$$

4. 设 X_1, X_2, X_3, X_4 是来自正态总体 $N(0, 3^2)$ 的随机样本, $X = a(X_1 - 2X_2)^2 + b(2X_3 - 3X_4)^2$ 。试确定 a 、 b 使统计量 X 服从 χ^2 分布, 并指出其自由度。

答案与提示: 依题意, 要使统计量 X 服从 χ^2 分布, 则必需使 $a^{1/2}(X_1 - 2X_2)$ 及 $b^{1/2}(2X_3 - 3X_4)$ 服从标准正态分布。解得

$$a = 1/45; \quad b = 1/117.$$

5. 设 X 和 Y 独立同分布 $N(0, 3^2)$, X_1, X_2, \dots, X_9 和 Y_1, Y_2, \dots, Y_9 分别是来自 X 和 Y 的简单抽样, 试确定统计量 $U = \frac{X_1 + \dots + X_9}{\sqrt{Y_1^2 + \dots + Y_9^2}}$ 所服从的分布。

答案与提示: 应用 t 分布的定义, 得

$$U = \frac{X_1 + \dots + X_9}{\sqrt{Y_1^2 + \dots + Y_9^2}} \sim t(9)$$

6. 设随机变量 $X \sim t(n)$ ($n > 1$), 试确定统计量 $Y = \frac{1}{X^2}$ 所服从的分布。

答案与提示: 先由 t 分布的定义知 $X = \frac{U}{\sqrt{V/n}}$, 再利用 F 分布的定义即可。

$$Y = \frac{1}{X^2} \sim F(n, 1)。$$

7. 设总体 X 服从正态分布 $N(0, 2^2)$ ，而 X_1, X_2, \dots, X_{15} 是来自总体 X 的简单随机样本，试确定随机变量 $Y = \frac{X_1^2 + \dots + X_{10}^2}{2(X_{11}^2 + \dots + X_{15}^2)}$ 所服从的分布。

答案与提示：由于 $(X_1^2/4 + \dots + X_{10}^2/4)/10 \sim \chi^2(10)$ ，

$$(X_{11}^2/4 + \dots + X_{15}^2/4)/5 \sim \chi^2(5)，$$

$$\text{故 } Y = \frac{X_1^2 + \dots + X_{10}^2}{2(X_{11}^2 + \dots + X_{15}^2)} \sim F(10, 5)$$

8. 设 X_1, X_2, \dots, X_n 为来自正态总体 $X \sim N(\mu, \sigma^2)$ 的一个样本， μ 已知，求 σ^2 的极大似然估计。

答案与提示：设 x_1, x_2, \dots, x_n 为样本 X_1, X_2, \dots, X_n 的一组观察值。则似然函数为

$$L(\mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}，$$

得 σ^2 的极大似然估计为

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2。$$

9. 设 $X \sim N(\mu, 1)$ ， X_1, X_2, \dots, X_n 为来自正态总体 X 的一个样本，试求 μ 的极大似然估计及矩估计。

答案与提示：矩估计法和极大似然估计法是点估计的两种常用方法，所谓矩估计法就是用样本的某种矩作为总体的相应矩的估计，因此需要首先计算（或已知）总体的某（几）种矩，由于本题只涉及一个未知参数，故只要知道总体的某一种矩即可。极大似然估计可依据四个步骤来完成，其关键是正确构造似然函数。

$$\mu \text{ 的极大似然估计为 } \hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i。$$

$$\mu \text{ 的矩估计为 } \hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i。$$

10. 设 X_1, X_2, \dots, X_n 为来自正态总体的一个样本，求下述各总体的密度函数中的未知参数的矩估计及极大似然估计。

$$(1) f(x, \theta) = \begin{cases} (\theta + 1)x^\theta, & 0 < x < 1, \\ 0, & \text{其它,} \end{cases}$$

其中 $\theta > -1$ 为未知参数。

$$(2) f(x) = \begin{cases} \frac{x}{\theta^2} e^{-x^2/2\theta^2}, & x > 0, \\ 0, & \text{其它}, \end{cases}$$

其中, $\theta > 0, \theta$ 为未知参数。

答案与提示：矩估计法和极大似然估计法是点估计的两种常用方法，所谓矩估计法就是用样本的某种矩作为总体的相应矩的估计，因此需要首先计算（或已知）总体的某（几）种矩，由于本例只涉及一个未知参数，故只要知道总体的某一种矩即可。极大似然估计可依据内容提要中的四个步骤来完成，其关键是正确构造似然函数。

$$(1) \text{矩估计: } \hat{\theta} = \frac{2\bar{X}-1}{1-\bar{X}}.$$

$$\text{极大似然估计: } \hat{\theta} = -1 - \frac{n}{\sum_{i=1}^n \ln X_i}.$$

$$(2) \text{矩估计: } \hat{\theta} = \bar{X} \sqrt{\frac{2}{\pi}}.$$

$$\text{极大似然估计: } \hat{\theta} = \sqrt{\frac{1}{2n} \sum_{i=1}^n x_i^2}$$

11. 设 X_1, X_2, \dots, X_n 为总体 X 的一个样本，且 X 服从几何分布，即

$$P\{X = k\} = (1-p)^{k-1} p, \quad k = 1, 2, 3, \dots,$$

求 p 的极大似然估计量。

答案与提示：极大似然估计为 $\hat{p} = 1/\bar{X}$

12. 设 X_1, X_2, \dots, X_n 为总体 X 的一个样本，且 X 服从参数为 m, p 的二项分布，求 p 的极大似然估计量。

答案与提示： p 的极大似然估计量为 $\hat{p} = \bar{X}/m$ 。

13. 设 X_1, X_2, \dots, X_n 为来自总体 X 的一个样本，且 $EX = \mu$ 存在，问统计量(1)(2)是否为 μ 的无偏估计。

$$(1) -4X_1 + 2X_2 + 5X_5;$$

$$(2) \frac{1}{10}(2X_1 + 3X_2 + X_{n-1} + 4X_n).$$

答案与提示：依据无偏估计定义， $-4X_1 + 2X_2 + 5X_5$ 不是 μ 的无偏估计；

$$\frac{1}{10}(2X_1 + 3X_2 + X_{n-1} + 4X_n) \text{ 是 } \mu \text{ 的无偏估计。}$$

14. 设总体 X 服从 $N(\mu, \sigma^2)$ ， X_1, X_2, X_3 为来自总体 X 的一个样本，试问统

计量 (1) (2) (3) 是否为 μ 的无偏估计, 并从无偏估计中找出较好的一个。

$$(1) \frac{1}{4}X_1 + \frac{1}{2}X_2 + \frac{1}{4}X_3; \quad (2) \frac{2}{3}X_1 + \frac{1}{4}X_2 + \frac{1}{12}X_3;$$

$$(3) \frac{1}{5}X_1 + \frac{3}{10}X_2 + \frac{1}{2}X_3。$$

答案与提示: 依据无偏估计定义, 统计量 (1) (2) (3) 均为 μ 的无偏估计。

由有效估计定义可判断 $\frac{1}{4}X_1 + \frac{1}{2}X_2 + \frac{1}{4}X_3$ 较好。

15. 设某种元件的使用寿命 X 的概率密度为

$$f(x; \theta) = \begin{cases} 2e^{-2(x-\theta)}, & x > \theta \\ 0, & x \leq \theta \end{cases},$$

其中 $\theta > 0$ 为未知参数。又设 x_1, x_2, \dots, x_n 是 X 的一组样本观察值, 求 θ 的极大似然估计值。

答案与提示: 构造似然函数 $L(\theta) = \prod_{i=1}^n 2e^{-2(x_i-\theta)} = 2^n e^{-2\sum_{i=1}^n (x_i-\theta)}$

$$\ln L = \ln 2^n - 2\sum_{i=1}^n (x_i - \theta)$$

$$\frac{d \ln L}{d\theta} = 2n \quad (\text{与参数 } \theta \text{ 无关})$$

由条件, 当 $x > \theta$ 时, $f(x) = 2e^{-2(x-\theta)}$ ($\theta > 0$), 所以当 $\theta = \min(x_1, x_2, \dots, x_n)$ 时, 似然函数 L 取得最大值, 从而知 $\hat{\theta} = \min(x_1, x_2, \dots, x_n)$ 。

16. 设总体 X 的概率分布为

X	0	1	2	3
P	θ^2	$2\theta(1-\theta)$	θ^2	$1-2\theta$

其中 θ ($0 < \theta < \frac{1}{2}$) 是未知参数, 利用总体 X 的如下样本值 3, 1, 3, 0, 3, 1, 2,

3, 求 θ 的矩估计值和极大似然估计值。

答案与提示: θ 的矩估计值为 $\hat{\theta} = 1/4$ 。

对于给定的样本值, 似然函数为 $L(\theta) = 4\theta^6(1-\theta)^2(1-2\theta)^4$, 解得

$$\theta_{1,2} = \frac{7 \pm \sqrt{13}}{12}$$

因 $\frac{7+\sqrt{13}}{12} > \frac{1}{2}$ 不合题意, 所以 θ 的极大似然估计值为

$$\hat{\theta} = \frac{7-\sqrt{13}}{12}。$$

17. 随机地从一批钉子中抽取 16 枚, 测得其长度 (单位 cm) 为

2.14, 2.10, 2.13, 2.15, 2.13, 2.12, 2.13, 2.10,

2.15, 2.12, 2.14, 2.10, 2.13, 2.11, 2.14, 2.11,

设钉长服从正态分布, 试就以下两种情况求总体均值 μ 的置信度为 90% 的置信区间:

(1) 若已知 $\sigma = 0.01$; (2) 若 σ 未知。

答案与提示: (1) μ 的置信度为 0.90 的置信区间是 (2.121, 2.129);

(2) μ 的置信度为 0.90 的置信区间是 (2.1175, 2.1325)。

18. 为了估计灯泡使用时数的均值 μ 及标准差 σ , 测试 10 只灯泡, 得 $\bar{x} = 1500$ 小时, $S = 20$ 。如果已知灯泡使用时数服从正态分布, 求总体均值 μ 、标准差 σ 的置信区间 (置信度为 0.95)。

答案与提示: (1) μ 的置信度为 0.90 的置信区间是 (1485.67, 1514.32);

(2) σ^2 的置信度为 0.95 的置信区间是 (189.47, 1333.33);

σ 的置信度为 0.95 的置信区间是 (13.76, 36.51)。

19. 随机的取某种炮弹 9 枚做试验, 得炮口速度的样本标准差 $S = 11$ (米/秒)。设炮口速度服从正态分布, 求这种炮弹炮口速度的标准差 σ 的 95% 的置信区间。

答案与提示: σ 的置信度 95% 的置信区间为 (7.43, 21.10)。

20. 随机的从 A 批导线中抽取 4 根, 从 B 批导线中抽取 5 根, 测得电阻 (Ω) 为

A 批导线: 0.143, 0.142, 0.143, 0.137,

B 批导线: 0.140, 0.142, 0.136, 0.138, 0.140,

设测定数据分别来自正态总体 $N(\mu_1, \sigma^2)$ 、 $N(\mu_2, \sigma^2)$, 且两样本相互独立。又 μ_1 、 μ_2 、 σ^2 均为未知, 试求 $\mu_1 - \mu_2$ 的置信度为 95% 的置信区间。

答案与提示: $\mu_A - \mu_B$ 的 0.95 的置信区间为 (-0.002, 0.006)。

21. 设两位化验员 A、B 独立地对某种聚合物含氮量用同样的方法各作 10 次测定, 其测定的样本方差依次为 $S_A^2 = 0.5419$, $S_B^2 = 0.6065$, 设 σ_A^2 、 σ_B^2 分别为 A、B 所测定的测定值总体的方差, 两总体均服从正态分布。试求方差比 σ_A^2 / σ_B^2 的置信度为 95% 的置信区间。

答案与提示: 方差比 σ_1^2 / σ_2^2 的置信度为 0.90 的置信区间是 (0.222, 3.601)。

22. 由经验知某零件重量 $X \sim N(15, 0.05)$ 。技术革新后, 抽了六个样品, 测得重量为 (单位: g) 14.7, 15.1, 14.8, 15.0, 15.2, 14.6, 已知方差不变, 问平均重量是否仍为 15? ($\alpha = 0.05$)

答案与提示: 依题意需检验假设 $H_0: \mu = 15$, 经计算知应接受 H_0 , 即认为平均重量仍是 15。

23. 原铸造成品率的平均值为 83.8%，今换用便宜的原料，成品率抽样数据（%）如下：83.9，84.6，82.4，84.1，84.9，82.9，85.2，83.3，82.0，83.5，问原料代用后，成品率是否发生了变化？（ $\alpha=0.05$ ）

答案与提示：依题意，可认为成品率这样的计量值数据服从正态分布，因此该问题即为方差未知的情况下，检验成品率的平均值是否仍为 83.8%。

检验结果：原料代用后，成品率无显著变化。

24. 设某产品的生产工艺发生了改变，在改变前后分别独立测了若干件产品的某项指标，其结果如下：

改变前：21.6，20.8，22.1，21.2，20.5，21.9，21.4；

改变后：24.1，23.8，24.7，24.0，23.7，24.3，24.5，23.9。

且假定产品的该项指标服从正态分布，求工艺改变前后该产品的此项指标稳定状况有无明显改变（ $\alpha=0.05$ ）？

答案与提示：依题意，设工艺改变后的总体为 $X \sim N(\mu_1, \sigma_1^2)$ ，工艺改变前的总体为 $Y \sim N(\mu_2, \sigma_2^2)$ ，从而问题化为检验假设 $H_{02}: \sigma_1^2 = \sigma_2^2$ 。

检验结果：认为工艺改变前后该产品的此项指标稳定状况无明显改变。

25. 机床厂某日从两台机器生产的同一零件中，分别抽取若干个样品测量的长度如下

第一台机器：6.2，5.7，6.5，6.0，6.3，5.8，5.7，6.0，6.0，5.8，6.0；

第二台机器：5.6，5.9，5.6，5.7，6.0，5.8，5.7，5.5，5.5。

问这两台机器的加工精度有无显著差异（ $\alpha=0.05$ ）？

答案与提示：依题意，可认为样品测量这样的计量值数据服从正态分布，因此比较两台机器的加工精度有无显著差异，即为检验假设 $H_0: \sigma_1^2 = \sigma_2^2$ 成立与否。

检验结果：认为两台机器的加工精度无显著差异。

26. 测得两批电子元件的样本的电阻（ Ω ）为

第一批：0.140，0.138，0.143，0.142，0.144，0.137，

第二批：0.135，0.140，0.142，0.136，0.138，0.140。

设两批电子元件的电阻分别服从正态总体 $N(\mu_1, \sigma^2)$ 、 $N(\mu_2, \sigma^2)$ ，且两样本相互独立。问这两批电子元件的电阻有无显著差异？（ $\alpha=0.05$ ）

答案与提示：显然该问题为在方差相等但未知的情况下，对两个正态总体均值是否相等的假设检验既要检验假设 $H_0: \mu_1 = \mu_2$ ， $H_1: \mu_1 \neq \mu_2$ 。

检验结果：认为两批电子元件电阻均值相等，无显著差异。

27. 假设 0.50，1.25，0.80，2.00 是来自总体 X 的简单随机样本值。已知 $Y = \ln x$ 服从正态分布 $N(\mu, 1)$ 。（1）求 X 的数学期望 EX （记 EX 为 b ）；（2）求 μ 的置信度

为 0.95 的置信区间 ; (3) 利用上述结果求 b 的置信度为 0.95 的置信区间。

答案与提示 :

$$(1) EX = e^{\mu + \frac{1}{2}}$$

(2) 由 $X : 0.50, 1.25, 0.80, 2.00$ 得

$$Y = \ln x : -0.693, 0.223, -0.223, 0.693.$$

从而得 μ 的置信度为 0.95 的置信区间为 $(-0.98, 0.98)$ 。

(3) 由上述结果知

$$P\{-0.98 < \mu < 0.98\} = 0.95$$

$$\text{即 } P\{e^{-0.98+0.5} < e^{\mu+0.5} < e^{0.98+0.5}\} = 0.95$$

$$\text{亦即 } P\{e^{-0.98+0.5} < b < e^{0.98+0.5}\} = P\{0.619 < b < 4.39\} = 0.95$$

故 b 的置信度为 0.95 的置信区间为 $(0.619, 4.39)$ 。

28. 设总体 X 服从正态分布 $N(\mu, \sigma^2)$ ($\sigma > 0$), 从总体中抽取简单随机样本 X_1, X_2, \dots, X_{2n} ($n > 2$), 其样本均值为 $\bar{X} = \frac{1}{2n} \sum_{i=1}^{2n} X_i$, 求统计量 $Y = \sum_{i=1}^n (X_i + X_{n+i} - 2\bar{X})^2$ 的数学期望 EY 。

解法 1 : 考虑 $(X_1 + X_{n+1}), (X_2 + X_{n+2}), \dots, (X_n + X_{2n})$, 将其视为取自总体 $N(2\mu, 2\sigma^2)$ 的简单随机样本, 则其样本均值、样本方差分别为

$$\frac{1}{n} \sum_{i=1}^n (X_i + X_{n+i}) = \frac{1}{n} \sum_{i=1}^{2n} X_i = 2\bar{X},$$

$$\frac{1}{n-1} \sum_{i=1}^n (X_i + X_{n+i} - 2\bar{X})^2 = \frac{1}{n-1} Y,$$

由于 $E(\frac{1}{n-1} Y) = 2\sigma^2$, 所以 $EY = (n-1)(2\sigma^2) = 2(n-1)\sigma^2$ 。

解法 2 : 记 $\bar{X}' = \frac{1}{n} \sum_{i=1}^n X_i$, $\bar{X}'' = \frac{1}{n} \sum_{i=1}^n X_{n+i}$, 显然有 $2\bar{X} = \bar{X}' + \bar{X}''$ 。因此

$$\begin{aligned} EY &= E\left[\sum_{i=1}^n (X_i + X_{n+i} - 2\bar{X})^2\right] = E\left\{\sum_{i=1}^n [(X_i - \bar{X}') + (X_{n+i} - \bar{X}'')]^2\right\} \\ &= E\left\{\sum_{i=1}^n [(X_i - \bar{X}')^2 + 2(X_i - \bar{X}')(X_{n+i} - \bar{X}'') + (X_{n+i} - \bar{X}'')^2]\right\} \\ &= E\left[\sum_{i=1}^n (X_i - \bar{X}')^2\right] + 0 + E\left[\sum_{i=1}^n (X_{n+i} - \bar{X}'')^2\right] \\ &= (n-1)\sigma^2 + (n-1)\sigma^2 = 2(n-1)\sigma^2 \end{aligned}$$

29. 设随机变量 X 的分布函数为 $F(x, \alpha, \beta) = \begin{cases} 1 - (\alpha/x)^\beta, & x > \alpha, \\ 0, & x \leq \alpha, \end{cases}$ 其中参数

$\alpha > 0, \beta > 1$ 。设 X_1, X_2, \dots, X_n 为来自总体 X 的简单随机样本，求：

- (1) 当 $\alpha = 1$ 时，求未知参数 β 的矩估计量；
- (2) 当 $\alpha = 1$ 时，求未知参数 β 的最大似然估计量；
- (3) 当 $\beta = 2$ 时，求未知参数 α 的最大似然估计量。

答案与提示：本题是一个常规题型，只要注意求连续型总体未知参数的矩估计和最大似然估计都须已知密度函数，从而先由分布函数求导得密度函数。

(1) 当 $\alpha = 1$ 时，参数 β 的矩估计量为 $\hat{\beta} = \frac{\bar{X}}{\bar{X} - 1}$ 。

(2) 当 $\alpha = 1$ 时，求未知参数 β 的最大似然估计量 $\hat{\beta} = \frac{n}{\sum_{i=1}^n \ln x_i}$ 。

(3) 当 $\beta = 2$ 时，求未知参数 α 的最大似然估计量， X 的概率密度为

$$f(x, \beta) = \begin{cases} \frac{2\alpha^2}{x^3}, & x > \alpha, \\ 0, & x \leq \alpha, \end{cases}$$

对于总体 X 的样本值 x_1, x_2, \dots, x_n ，似然函数为

$$L(\alpha) = \prod_{i=1}^n f(x_i; \alpha) = \begin{cases} \frac{2^n \alpha^{2n}}{(x_1 x_2 \cdots x_n)^3}, & x_i > \alpha (i=1, 2, \dots, n), \\ 0, & \text{其他.} \end{cases}$$

当 $x_i > \alpha (i=1, 2, \dots, n)$ 时， α 越大， $L(\alpha)$ 越大，即 α 的最大似然估计值为

$$\hat{\alpha} = \min\{x_1, x_2, \dots, x_n\},$$

于是 α 的最大似然估计量为

$$\hat{\alpha} = \min\{X_1, X_2, \dots, X_n\}。$$