# 概率论与数理统计 B 习题五答案

## A

### 第五章 中心极限定理：

1. 在人寿保险公司里有3000个同龄的人参加人寿保险。在1年内每人的死亡率为0.1%，参加保险的人在 1 年的第一天交付保险费 10 元，死亡时家属可以从保险公司领取 2000 元。试用中心极限定理求保险公司亏本的概率。

解：设死亡人数为$X, X\sim B(3000,0.001)$，保险公司亏本当且仅当$2000X>10\times 3000$，即$X>15$。于是，由棣莫弗—拉普拉斯定理，公司亏本的概率为

$$\begin{aligned}P(X>15)&=P\left(\frac{X-np}{\sqrt{np(1-p)}}>\frac{15-np}{\sqrt{np(1-p)}}\right)\\&=p\left(\frac{x-3}{3\times 0.999}>\frac{15-3}{1.73}\right)\\&\approx 1-\Phi(6.93)=0\end{aligned}$$

2. 一保险公司有 1 万个投保人，每个投保人的索赔金额的数学期望为 250 元，标准差为 500，求索赔金额不超过 260 万元的概率。

解：设第$i$个投保人的索赔金额为随机变量$X_i\ (i=1,2,\cdots,10000)$，则$X_1,X_2,\cdots,X_{10000}$

独立同分布，且$E(X_i)=250$，$D(X_i)=500^2\quad (i=1,2,\cdots,10000)$，

索赔总金额不超过 2600000 元可表示为事件$\{\sum_{i=1}^{10000}X_i\le 2600000\}$，由中心极限定理有

$$P\{\sum_{i=1}^{10000}X_i\le 2600000\}=\Phi(\frac{2600000-10000\times 250}{\sqrt{10000}\times 500})=\Phi(2)=0.9772\text{ 。}$$

3. 某单位设置一电话总机，共有 200 个电话分机，若每个分机有 5%的时间要使用外线通话，假设每个分机是否使用外线通话是相互独立的，问总机要有多少条外线才能保证每个分机正常使用外线的概率不小于 90%?

解：设$X$为 200 个电话分机中要使用外线通话的分机数，则$X\sim B(200,0.05)$，如果

有外线$n$条，则$P\{X\le n\}\ge 0.9$，由中心极限定理得：

$$P\{X\le n\}=\Phi(\frac{n-200\times 0.05}{\sqrt{200\times 0.05\times 0.95}})=\Phi(\frac{n-10}{\sqrt{9.5}})\ge 0,$$

得 $\Phi(1.28)\geq 0.90$，所以 $\dfrac{n-10}{\sqrt{9.5}}\geq 1.28$，解得 $n\geq 13.945$，从而得 $n=14$。

4. 设某电话总机要为 2000 个用户服务，在最忙时，平均每户有 3% 的时间占线，假设各户是否打电话是相互独立的，试求若想以 99% 的可能性满足用户的要求，最少需要设多少条线路？

解：设电话交换台每小时呼叫次数为 $X$，在每小时每户用线的概率 $P=0.03$，由泊松分布近似可取 $\lambda=np=2000\times 0.03=60$，因此，$X$ 服从参数 $\lambda=60$ 的泊松分布。

设要求的最小线路数为 $m$：$P\{0\leq X\leq m\}>0.99$

$$E(X)=\lambda=60,\quad D(X)=\lambda=60$$

由泊松分布的正态逼近可得：

$$P\{0\leq X\leq m\}=\frac{1}{\sqrt{2\pi}}\int_{\frac{0-60}{\sqrt{60}}}^{\frac{m-60}{\sqrt{60}}}e^{-\frac{t^2}{2}}dt=\Phi(\frac{m-60}{\sqrt{60}})-\Phi(-\frac{60}{\sqrt{60}})\approx\Phi(\frac{m-60}{\sqrt{60}})$$

根据题意 $\Phi(\dfrac{m-60}{\sqrt{60}})>0.99$，可得：

$$\frac{m-60}{\sqrt{60}}>2.327\text{，从而 } m\geq 60+2.237\times 7.746\approx 78.023$$

所以最少需要 79 条线路。

5. 某供电站供应某地区 1000 户居民用电，各户用电情况相互独立。已知每户每日用电量（单位：度）在[0，20]上均匀分布。问：供电站每天至少向该地区供应多少度电才能以 0.99 的概率保证该地区居民供应电量的需求？

解：用 $X_i$ 表示居民每户每日的用电量，则 $X_i$ 的密度函数为：

$$f_{X_i}(x)=\begin{cases}\dfrac{1}{20}, & 0\leq x\leq 20\\ 0, & \text{其他}\end{cases}$$

从而：

$$E(X)=\frac{20+0}{2}=10,\quad D(X)=\frac{(20-0)^2}{12}=\frac{100}{3}$$

令 X 表示该地区居民的用电量，则

$$X=\sum_{k=1}^{1000}X_k$$

所以：

$$E(X)=E\left(\sum_{k=1}^{1000}X_k\right)=10000,\quad D(X)=D\left(\sum_{k=1}^{1000}X_k\right)=\frac{100000}{3}$$

于是若设供电站每天至少应供应 $n$ 度电，则

$$0.99 \le P(X \le n) = P\left(\frac{X-E(X)}{\sqrt{D(X)}} \le \frac{n-E(X)}{\sqrt{D(X)}}\right) = P\left(\frac{X-10000}{\sqrt{100000/3}} \le \frac{n-1000}{\sqrt{100000/3}}\right)$$

$$\approx \Phi\left(\frac{n-1000}{\sqrt{100000/3}}\right) = \Phi\left(\frac{n-1000}{182.6}\right)$$

即：$\frac{n-1000}{182.6} \ge 2.33$， $n \ge 10425.4$

6. 机器包装某种面包时，每袋面包的净重为随机变量，平均重量为 100 克，标准差为 10 克。一箱内装 200 袋面包，求一箱面包的净重大于 20500 克的概率。

解：设箱中第 $i$ 袋面包的净重为 $X_i$，则 $X_i$ 独立同分布，

且 $E(Xi)=100$，$\mathrm{Var}(Xi)=100$，

由中心极限定理得，所求概率为：

$$P\left(\sum_{i=1}^{200} X_i > 20500\right) \approx 1-\Phi\left(\frac{20500-200\times 100}{\sqrt{200\times 100}}\right)$$

$$=1-\Phi(3.54)$$

$$=0.0002$$

故一箱味精的净重大于 20500 克的概率为 0.0002.

## 第六章 样本及抽样分布：

7. 从正态总体 $N(4.2,25)$ 中抽取容量为 $n$ 的样本，若要求其样本均值位于区间（2.2，6.2）内的概率不小于 0.95，则样本容量 $n$ 至少应该取多大？

解：由于

$$\frac{\overline{X}-4.2}{\sqrt{25/n}} = \frac{\overline{X}-4.2}{5}\sqrt{n} \sim N(0,1)$$

所以：

$$0.95 \le P(2.2 < \overline{X} \le 6.2) = P\left(\frac{2.2-4.2}{5}\sqrt{n} < \frac{\overline{X}-4.2}{5}\sqrt{n} \le \frac{6.2-4.2}{5}\sqrt{n}\right)$$

$$=\Phi\left(\frac{6.2-4.2}{5}\sqrt{n}\right)-\Phi\left(\frac{2.2-4.2}{5}\sqrt{n}\right)=\Phi\left(\frac{2}{5}\sqrt{n}\right)-\Phi\left(\frac{-2}{5}\sqrt{n}\right)=2\Phi\left(\frac{2}{5}\sqrt{n}\right)-1$$

即：

$$2\Phi\left(\frac{2}{5}\sqrt{n}\right)-1 \ge 0.95，\text{即}\ \Phi\left(\frac{2}{5}\sqrt{n}\right) \ge 0.975 = \Phi(1.96)$$

由分布函数的单调性有：

$$\frac{2}{5}\sqrt{n} \ge 1.96, n \ge (2.5\times 1.96)^2 = 24.01$$

可见，样本容量至少取 25。

8. 设总体 $X \sim N(\mu,4)$，有样本 $X_1,X_2,\cdots,X_n$，求当样本容量 $n$ 为多大时，$P\{|\overline{X}-\mu| \le 0.1\} = 0.95$。

解：因为 $\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}\sim N(0,1)$，所以

$$P\{|\overline{X}-\mu|\le 0.1\}=P\left\{\frac{-0.1}{\frac{2}{\sqrt{n}}}\le\frac{\overline{X}-\mu}{\frac{2}{\sqrt{n}}}\le\frac{0.1}{\frac{2}{\sqrt{n}}}\right\}=\Phi(0.05\sqrt{n})-\Phi(-0.05\sqrt{n})=2\Phi(0.05\sqrt{n})-1$$，得：

$2\Phi(0.05\sqrt{n})-1=0.95$， $\Phi(0.05\sqrt{n})=0.975$， 由 $\Phi(1.96)=0.975$， 可得 $0.05\sqrt{n}=1.96$，

于是得 $n=1536.6$， 即 $n=1537$ 。

9. 设 $X_1,X_2,\cdots,X_5$ 是独立且服从相同分布的随机变量，且每一个 $X_i(i=1,2,\cdots,5)$ 都服从 $N(0,1)$。（1）试给出常数 $c$，使得 $c(X_1^2+X_2^2)$ 服从 $\chi^2$ 分布，并指出它的自由度；（2）试给出常数 $d$，使得 $d\dfrac{X_1+X_2}{\sqrt{X_3^2+X_4^2+X_5^2}}$ 服从 $t$ 分布，并指出它的自由度。

解：（1）易见，$X_1^2+X_2^2$ 即为二个独立的服从 $N(0,1)$ 的随机变量平方和，服从 $\chi^2(2)$ 分布，即 $c=1$；自由度为 2。

（2）由于 $X_1+X_2\sim N(0,2)$，则 $\dfrac{X_1+X_2}{\sqrt{2}}\sim N(0,1)$，又 $X_3^2+X_4^2+X_5^2\sim\chi^2(3)$，$\dfrac{X_1+X_2}{\sqrt{2}}$ 与 $X_3^2+X_4^2+X_5^2$ 相互独立，则

$$\frac{(X_1+X_2)/\sqrt{2}}{\sqrt{(X_3^2+X_4^2+X_5^2)/3}}\sim t(3)，\quad 即\ \frac{\sqrt{6}}{2}\frac{X_1+X_2}{\sqrt{X_3^2+X_4^2+X_5^2}}\sim t(3)，$$

即 $d=\dfrac{\sqrt{6}}{2}$，自由度为 3。

## 第七章 点估计：

10. 设 $X_1,X_2,\cdots,X_n$ 是取自总体 $X$ 的一个样本，在下列情形下，试求总体参数的矩估计与最大似然估计：（1）$X\sim B(n,p)$，其中 $p$ 未知，$0<p<1$；（2）$X\sim E(\lambda)$，其中 $\lambda$ 未知，$\lambda>0$。

解：（1）$E(X)=p$，故 $p$ 的矩估计量有 $\hat{p}=\overline{X}$ 。

另，$X$ 的分布律为 $P(X=x)=p^x(1-p)^{1-x},x=0,1$，故似然函数为

$$L(p)=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}$$

对数似然函数为：

$$\ln L(p)=\left(\sum_{i=1}^{n}X_i\right)\ln p+\left(n-\sum_{i=1}^{n}X_i\right)\ln(1-p)$$

令 $\frac{d\ln L(p)}{dp}=\frac{\sum_{i=1}^{n}X_i}{p}-\frac{n-\sum_{i=1}^{n}X_i}{1-p}=0$，解得 $p$ 的最大似然估计量 $\hat{p}=\frac{1}{n}\sum_{i=1}^{n}X_i=\overline{X}$ 。

可以看出 $p$ 的矩估计量与最大似然估计量是相同的。

（2）$E(X)=\frac{1}{\lambda}$，令 $\frac{1}{\lambda}=\overline{X}$，故 $\lambda$ 的矩估计量 $\hat{\lambda}=\frac{1}{\overline{X}}$ 。

另，X 的密度函数为

$$f_X(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

故似然函数为

$$L(\lambda)=\begin{cases}\lambda^n e^{-\lambda\sum_{i=1}^{n}X_i} & X_i>0, i=1,2,\cdots,n\\ 0 & \text{其他}\end{cases}$$

对数似然函数为

$$\ln L(\lambda)=n\ln\lambda-\lambda\sum_{i=1}^{n}X_i$$

$$\frac{d\ln L(\lambda)}{d\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}X_i=0$$

解得 $\lambda$ 的最大似然估计量 $\hat{\lambda}=\frac{n}{\sum_{i=1}^{n}X_i}=\frac{1}{\overline{X}}$ 。

可以看出 $\lambda$ 的矩估计量与最大似然估计量是相同的。

11. 设 $X_1,X_2,\cdots,X_n$ 是取自总体 $X$ 的一个样本，其中 $X$ 服从参数为 $\lambda$ 的泊松分布，其中 $\lambda$ 未知，$\lambda>0$，求 $\lambda$ 的矩估计与最大似然估计，如得到一组样本观测值：

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 频数 | 17 | 20 | 10 | 2 | 1 |

求 $\lambda$ 的矩估计值与最大似然估计值。

解：$E(X)=\lambda$，故 $\lambda$ 的矩估计量 $\hat{\lambda}=\overline{X}$ 。

由样本观测值可算得

$$\overline{X}=\frac{0\times17+1\times20+2\times10+3\times2+4\times1}{50}=1,$$

另，$X$ 的分布律为

$$P(X=x)=e^{-\lambda}\frac{\lambda^x}{x!},x=0,1,2,\cdots$$

故似然函数为

$$L(\lambda)=e^{-n\lambda}\frac{\lambda\sum_{i=1}^{n}X_i}{X_1!\cdots X_{n!}},X_i=0,1,2,\cdots,i=1,2,\cdots,n$$

对数似然函数为

$$\ln L(\lambda)=-n\lambda+\left(\sum_{i=1}^{n}X_i\right)\ln\lambda-\sum_{i=1}^{n}\ln\left(X_i!\right)$$

$$\frac{d\ln L(\lambda)}{d\lambda}=-n+\frac{\sum_{i=1}^{n}X_i}{\lambda}=0$$

解得$\lambda$的最大似然估计量$\hat{\lambda}=\frac{\sum_{i=1}^{n}X_i}{n}=\overline{X}$，故$\lambda$的最大似然估计值$\hat{\lambda}=1$。

12. 设$X_1,X_2,\cdots,X_n$是取自总体$X$的一个样本，其中$X$服从区间$(0,\theta)$的均匀分布，其中$\theta>0$未知，求$\theta$的矩估计，且$\theta$的矩估计是否是$\theta$的无偏估计？

解：$E(X)=\frac{\theta}{2}$，令$\frac{\theta}{2}=\overline{X}$，故$\theta$的矩估计量$\hat{\theta}=2\overline{X}$。

$$E(\hat{\theta})=E(2\overline{X})=2E(\overline{X})=2E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{2}{n}\sum_{i=1}^{n}E(X_i)=\frac{2}{n}\sum_{i=1}^{n}\frac{\theta}{2}=\theta,$$

故$\theta$的矩估计量$2\overline{X}$是$\theta$的无偏估计。

13. 设$X_1,X_2,\cdots,X_n$是取自总体$X$的一个样本，$X$的密度函数为

$$f(x)=\begin{cases}\frac{2x}{\theta^2} & 0<x<\theta\\ 0 & \text{其他}\end{cases}$$

其中$\theta>0$未知，求$\theta$的矩估计。

解：$E(X)=\int_0^{\theta}x\cdot\frac{2x}{\theta^2}dx=\frac{2}{3}\theta$，令$\frac{2}{3}\theta=\overline{X}$，故$\theta$的矩估计量为$\hat{\theta}=\frac{3}{2}\overline{X}$。

14. 设$X_1,X_2,\cdots,X_n$是取自总体$X$的一个样本，$X$的密度函数为

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

其中$\theta>0$未知，求$\theta$的矩估计和最大似然估计。

解：$E(X)=\int_0^{1}x\cdot(\theta+1)x^{\theta}dx=\frac{\theta+1}{\theta+2}$，令$\frac{\theta+1}{\theta+2}=\overline{X}$，故$\theta$的矩估计量为$\hat{\theta}=\frac{1-2\overline{X}}{\overline{X}-1}$，另，似然函数

$$L(\theta)=\begin{cases}(\theta+1)^n\prod_{i=1}^{n}X_i^{\theta} & 0<X_i<1\\ 0 & \text{其他}\end{cases}$$

对数似然函数为

$$\ln L(\theta)=n\ln(\theta+1)+\theta\sum_{i=1}^{n}\ln X_i,$$

$$\frac{d\ln L(\theta)}{d\theta}=\frac{n}{\theta+1}+\sum_{i=1}^{n}\ln X_i=0$$

解得$\theta$的最大似然估计量为$\hat{\theta}=-1-\dfrac{n}{\sum\limits_{i=1}^{n}\ln X_i}$。

15. 设$X_1,X_2,\cdots,X_n$是取自总体$X$的一个样本，总体$X$服从参数为$p$的几何分布，即

$P(X=x)=p(1-p)^{x-1},(x=1,2,3,\cdots)$，其中$p$未知，$0<p<1$，求$p$的最大似然估计。

解：似然函数$L(p)=p^n(1-p)^{\sum_{i=1}^{n}X_i-n}$，对数似然函数：

$$\ln L(p)=n\ln p+\left(\sum_{i=1}^{n}X_i-n\right)\ln(1-p)$$

$$\frac{d\ln L(p)}{dp}=\frac{n}{p}-\frac{\sum\limits_{i=1}^{n}X_i-n}{1-p}=0$$

解得$p$的最大似然估计量为$\hat{p}=\dfrac{1}{\overline{X}}$。

16. 已知某路口车辆经过的时间间隔服从指数分布$E(\lambda)$，其中$\lambda>0$未知，现在观测到六个时间间隔数据（单位：s）：1.8、3.2、4、8、4.5、2.5，试求该路口车辆经过的平均时间间隔的矩估计值与最大似然估计值。

解：根据习题 11 的结果，$\lambda$的矩估计和最大似然估计量都为$\dfrac{1}{\overline{X}}$，故平均时间间隔的矩估计和最大似然估计都为$\dfrac{1}{\hat{\lambda}}$，即为$\overline{X}$。

由样本观测值可算得$\overline{X}=\dfrac{1}{6}\times(1.8+3.2+4+8+4.5+2.5)=4$。

17. 设$X_1,X_2,X_3$为总体$X\sim N(\mu,\sigma^2)$的样本，证明

$$\hat{\mu}_1=\frac{1}{6}X_1+\frac{1}{3}X_2+\frac{1}{2}X_3,\hat{\mu}_2=\frac{2}{5}X_1+\frac{1}{5}X_2+\frac{2}{5}X_3$$

都是总体均值$\mu$的无偏估计，并进一步判断哪一个估计有效。

证明：$E(\hat{\mu}_1)=E\left(\dfrac{1}{6}X_1+\dfrac{1}{3}X_2+\dfrac{1}{2}X_3\right)$

$$=\frac{1}{6}E(X_1)+\frac{1}{3}E(X_2)+\frac{1}{2}E(X_3)=\left(\frac{1}{6}+\frac{1}{3}+\frac{1}{2}\right)E(X)=E(X)=\mu,$$

$$\begin{aligned}E(\hat{\mu}_2)&=E\left(\frac{2}{5}X_1+\frac{1}{5}X_2+\frac{2}{5}X_3\right)\\&=\frac{2}{5}E(X_1)+\frac{1}{5}E(X_2)+\frac{2}{5}E(X_3)=\left(\frac{2}{5}+\frac{1}{5}+\frac{2}{5}\right)E(X)=E(X)=\mu\end{aligned},$$

所以$\hat{\mu}_1,\hat{\mu}_2$都是总体均值$\mu$的无偏估计，又

$$D(\hat{\mu}_1)=D\left(\frac{X_1}{6}+\frac{X_2}{3}+\frac{X_3}{2}\right)$$
$$=\frac{1}{36}D(X_1)+\frac{1}{9}D(X_2)+\frac{1}{4}D(X_3)=\left(\frac{1}{36}+\frac{1}{9}+\frac{1}{4}\right)D(X)=\frac{7}{18}D(X)=\frac{7}{18}\sigma^2,$$

$$D(\hat{\mu}_2)=D\left(\frac{2}{5}X_1+\frac{1}{5}X_2+\frac{2}{5}X_3\right)$$
$$=\frac{4}{25}D(X_1)+\frac{1}{25}D(X_2)+\frac{4}{25}D(X_3)=\frac{9}{25}D(X)=\frac{9}{25}\sigma^2,$$

可见 $D(\hat{\mu}_2)<D(\hat{\mu}_1)$，所以二个估计量中 $\hat{\mu}_2$ 更有效。

# B

## 第五章 中心极限定理：

1. 某地进行的抽样调查结果显示，考生的外语成绩（百分制）近似的服从正态分布，平均成绩为 72 分，96 分以上的占考生总数的 2.3%。试求考生的外语成绩在 60 分至 84 分之间的概率。

**解** 以 $X$ 记考生的外语成绩，则 $X\sim N(72,\sigma^2)$. 又由题设知

$$P\{X\geqslant 96\}=P\left\{\frac{X-72}{\sigma}\geqslant\frac{96-72}{\sigma}\right\}=1-\Phi\left(\frac{24}{\sigma}\right)=0.023,$$

即 $\Phi\left(\frac{24}{\sigma}\right)=0.977$. 查正态分布表知，$24/\sigma=2$，解得 $\sigma=12$. 所以，$X\sim N(72,12^2)$. 于是

$$P\{60\leqslant X\leqslant 84\}=P\left\{\frac{60-72}{12}\leqslant X\leqslant\frac{84-72}{12}\right\}$$
$$=\Phi(1)-\Phi(-1)=2\Phi(1)-1=0.682.$$

2. 一生产线生产的产品成箱包装，每箱的重量是随机的。假设每箱平均重 50 千克，标准差为 5 千克。若用最大载重量为 5 吨的汽车承运，试利用中心极限定理说明每辆车最多装多少箱，才能保证不超载的概率大于 0.977（$\Phi(2)=0.977$，其中 $\Phi(z)$ 是标准正态分布函数）。

解：设第 $i$ 箱的重量为 $X_i$， $i=1,2,\cdots,n$， $Y_n=\sum_{i=1}^{n}X_i$ 为 $n$ 箱的重量，由题设知：

$E(X_i)=50$， $\sqrt{D(X_i)}=5$， $i=1,2,\cdots,n$， $E(Y)=50n$， $\sqrt{D(Y)}=5\sqrt{n}$

且随机变量 $X_1,X_2,\cdots,X_n$ 相互独立且同分布，于是由独立且同分布中心极限定理，

$Y_n$ 近似服从正态分布 $N(50n,25n)$，若要求 $P\{Y_n=\sum_{i=1}^{n}X_i\le 5000\}\ge 0.977$，

即 $0.977\le P\{Y_n=\sum_{i=1}^{n}X_i\le 5000\}\approx\Phi\left(\frac{5000-50n}{5\sqrt{n}}\right)$，于是应有 $\frac{5000-50n}{5\sqrt{n}}\ge 2$，

$\frac{1000-10n}{\sqrt{n}} \geq 2$，从而有 $n \leq 98.0199$，即最多可以装 98 箱。

3. 某仪器上的一个易损元件坏了，现买回 80 个这种元件，更换一个后，其余作后备用，以便再损坏时能够立即更换。已知买回的这批元件中每一个的使用寿命（以 $h$ 计）都服从指数分布 $Z(0.2)$，试求这批元件使用的总时数能超过 500$h$ 的概率。（$\Phi(2.24)=0.9875$）

解：$X_i$ 表示第 $i$ 个元件的寿命，$X_i$ i.i.d. $Z(0.2)$，$i=1,2,...,80$，则 $E(X_i)=5$，

$D(X_i)=25$，故 $X=\sum_{i=1}^{80} X_i$ 表示 80 个元件的总寿命，且 $E(X)=\sum_{i=1}^{80} E(X_i)=400$，

$D(X)=\sum_{i=1}^{80} D(X_i)=2000$，

根据独立同分布中心极限定理，有

$$P(X>500)=1-P(X\leq 500)\approx 1-\Phi\left(\frac{500-400}{\sqrt{2000}}\right)=1-\Phi(2.24)=0.0125。$$

4. 从次品率为 0.05 的一批产品中随机地取 200 件产品，试求取出的产品中至少有 3 个次品的概率。

解：设 $X$ 表示取出的 200 件产品中的次品数的数学期望和方差为：

$$E(X)=np=200\times 0.05=10，\quad D(X)=np(1-p)=200\times 0.05\times(1-0.05)=9.5，$$

由棣莫弗—拉普拉斯中心极限定理得：

$$P\{X\geq 3\}=1-P\{0\leq X<3\}\approx 1-\Phi(\frac{3-10}{\sqrt{9.5}})+\Phi(\frac{0-10}{\sqrt{9.5}})=1-\Phi(-2.27)+\Phi(-3.24)\approx 0.989。$$

5. 设某种电气元件不能承受超负荷试验的概率为 0.05．现在对 100 个这样独立工作的元件进行超负荷试验，以 $X$ 表示不能承受试验而烧毁的元件数，请利用中心极限定理计算 $P(5\leq X\leq 10)$。

解：根据题意知，不能承受试验而烧毁的元件数 $X\sim B(n,p)$，

根据棣莫弗-拉普拉斯定理，$X$ 近似服从正态分布 $N(np,npq)$，

其中 $n$=100，$p$=0.05，$q$=0.95：

$$\mathbf{P}\{5\leq X\leq 10\}=\mathbf{P}\left\{\frac{5-np}{\sqrt{npq}}\leq\frac{X-np}{\sqrt{npq}}\leq\frac{10-np}{\sqrt{npq}}\right\}=\mathbf{P}\left\{0\leq\frac{X-5}{\sqrt{4.75}}\leq\frac{10-5}{\sqrt{4.75}}\right\}$$

$$=\mathbf{P}\left\{0\leq\frac{X-5}{\sqrt{4.75}}\leq 2.29\right\}\approx\Phi(2.29)-\Phi(0)=\Phi(2.29)\text{-}0.5.$$

6. 假设一条生产线生产的产品合格率是 0.8.要使一批产品的合格率达到在 76%与 84%之间的概率不小于 90%，问这批产品至少要生产多少件？

解：令 $X_i=\begin{cases}1,\text{若第}i\text{个产品是合格品,}\\0,\text{其他情形.}\end{cases}$

而至少要生产 $n$ 件，则 $i$=1,2,…,$n$,且 $X_1$，$X_2$，…，$X_n$ 独立同分布，$p=P\{X_i=1\}$=0.8.现要求 $n$,使得

$$P\{0.76 \le \frac{\sum_{i=1}^{n} X_i}{n} \le 0.84\} \ge 0.9.$$

即

$$P\{\frac{0.76n-0.8n}{\sqrt{n\times 0.8\times 0.2}} \le \frac{\sum_{i=1}^{n} X_i - 0.8n}{\sqrt{n\times 0.8\times 0.2}} \le \frac{0.84n-0.8n}{\sqrt{n\times 0.8\times 0.2}}\} \ge 0.9$$

由中心极限定理得

$$\Phi\left(\frac{0.84n-0.8n}{\sqrt{0.16n}}\right) - \Phi\left(\frac{0.76n-0.8n}{\sqrt{0.16n}}\right) \ge 0.9,$$

整理得 $\Phi\left(\frac{\sqrt{n}}{10}\right) \ge 0.95,$ 查表 $\frac{\sqrt{n}}{10} \ge 1.645,$ $n\ge 270.60,$ 故取 $n=271$。

## 第六章 样本及抽样分布：

7. 设总体 $X \sim N(0,1^2)$，从总体中取一个容量为 6 的样本 $X_1$，$X_2$，$\cdots$，$X_6$，设 $Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$，试确定常数 $C$，使随机变量 $CY$ 服从 $\chi^2$ 分布。

解：因为 $X_1+X_2+X_3 \sim N(0,\sqrt{3}^2)$，所以 $\frac{X_1+X_2+X_3}{\sqrt{3}} \sim N(0,1^2)$。

于是 $(\frac{X_1+X_2+X_3}{\sqrt{3}})^2 \sim \chi^2(1)$，同理 $(\frac{X_4+X_5+X_6}{\sqrt{3}})^2 \sim \chi^2(1)$。

由 $\chi^2$ 分布的可知性，故 $\frac{1}{3}Y=(\frac{X_1+X_2+X_3}{\sqrt{3}})^2+(\frac{X_4+X_5+X_6}{\sqrt{3}})^2 \sim \chi^2(2)$

可知 $C=\frac{1}{3}$。

8. 设总体 $\boldsymbol{X} \sim \boldsymbol{N}(0,4)$，$\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3,\boldsymbol{X}_4,\boldsymbol{X}_5$ 是来自 $X$ 的简单随机样本。记 $Z=C\frac{X_1+X_2}{\sqrt{X_3^2+X_4^2+X_5^2}}$，试确定常数 $C$，使 $Z$ 服从 $t$ 分布并确定自由度。

解：$\boldsymbol{X}_1+\boldsymbol{X}_2 \sim \boldsymbol{N}(0,8) \Rightarrow \frac{\boldsymbol{X}_1+\boldsymbol{X}_2}{\sqrt{8}} \sim \boldsymbol{N}(0,1)$

$$\frac{\boldsymbol{X}_3}{2},\frac{\boldsymbol{X}_4}{2},\frac{\boldsymbol{X}_5}{2}\,\text{i.i.d.}\boldsymbol{N}(0,1) \Rightarrow \left(\frac{\boldsymbol{X}_3}{2}\right)^2+\left(\frac{\boldsymbol{X}_4}{2}\right)^2+\left(\frac{\boldsymbol{X}_5}{2}\right)^2=\frac{\boldsymbol{X}_3^2+\boldsymbol{X}_4^2+\boldsymbol{X}_5^2}{4} \sim \boldsymbol{\chi}^2(3)$$

$$\frac{\frac{\boldsymbol{X}_1+\boldsymbol{X}_2}{\sqrt{8}}}{\sqrt{\frac{\left(\boldsymbol{X}_3^2+\boldsymbol{X}_4^2+\boldsymbol{X}_5^2\right)/4}{3}}}=\sqrt{\frac{3}{2}}\frac{\boldsymbol{X}_1+\boldsymbol{X}_2}{\sqrt{\boldsymbol{X}_3^2+\boldsymbol{X}_4^2+\boldsymbol{X}_5^2}} \sim \boldsymbol{t}(3) \Rightarrow \boldsymbol{C}=\sqrt{\frac{3}{2}}$$，自由度为 3

9. 设 $X_1,X_2,\cdots,X_n$ $(n\ge 2)$ 为来自总体 $N(0,\sigma^2)$ 的简单随机样本，$\bar{X}$ 为样本均值，$S^2$

为样本方差，证明：$\dfrac{(X_1+X_2)^2}{(X_3-X_4)^2} \sim F(1,1)$。

证明：

$$X_1+X_2 \sim N(0,2\sigma^2), X_3-X_4 \sim N(0,2\sigma^2)$$

$$\frac{X_1+X_2}{\sqrt{2}\sigma} \sim N(0,1) \quad \frac{X_3-X_4}{\sqrt{2}\sigma} \sim N(0,1),$$

由定理3.4知

$$\left[\frac{X_1+X_2}{\sqrt{2}\sigma}\right]^2 \sim x^2(1) \quad \left[\frac{X_3-X_4}{\sqrt{2}\sigma}\right]^2 \sim x^2(1),$$

所以 $X_1+X_2$ 与 $X_3-X_4$ 相互独立，

因此：$\dfrac{(X_1+X_2)^2}{(X_3-X_4)^2} \sim F(1,1)$。

10. 从正态总体 $X \sim N(34,6^2)$ 中抽取容量为 $n$ 的样本，如果要求其样本均值位于区间 $(29,39)$ 内的概率不小于 $0.95$，试问样本容量至少应取多大？

解：因为 $X \sim N(34,6^2)$，所以 $\bar{X} \sim N(34,6^2/n)$，欲使

$$0.95 \le P\left\{29 < \bar{X} < 39\right\} = \Phi\left(\frac{39-34}{6/\sqrt{n}}\right) - \Phi\left(\frac{29-34}{6/\sqrt{n}}\right) = \Phi\left(\frac{5\sqrt{n}}{6}\right) - \Phi\left(\frac{-5\sqrt{n}}{6}\right),$$

即 $2\Phi\left(\dfrac{5\sqrt{n}}{6}\right)-1 \ge 0.95$，$\Phi\left(\dfrac{5\sqrt{n}}{6}\right) \ge 0.975$，$\dfrac{5\sqrt{n}}{6} \ge 1.96$，$n \ge \left(\dfrac{6\times 1.96}{5}\right)^2 = 5.53 \approx 6$，

即样本容量至少应取 $n \ge 6$。

11. 在天平上重复称一重为 $\alpha$ 的物品，假设各次称量结果相互独立且同时服从正态分布 $N(\alpha,0.2^2)$。以 $\bar{X}_n$ 表示 $n$ 次称量结果的算术平均值，则为使 $P\{|\bar{X}_n-\alpha|<0.1\} \ge 0.95$，试求 $n$ 的最小值。

解：设第 $i$ 次称量结果为 $X_i$，$i=1,2,\cdots,n$，由题设知，$X_1,X_2,\cdots,X_n$ 相互独立且同时服从正态分布 $N(\alpha,0.2^2)$，所以其算术平均值 $\bar{X}_n \sim N(\alpha,\dfrac{0.2^2}{n})$，于是

$$P\{|\bar{X}_n-\alpha|<0.1\} = P\{\frac{|\bar{X}_n-\alpha|}{0.2/\sqrt{n}} < \frac{0.1}{0.2/\sqrt{n}}\}$$

$$=\Phi\left(\frac{\sqrt{n}}{2}\right)-\Phi\left(-\frac{\sqrt{n}}{2}\right)=2\Phi\left(\frac{\sqrt{n}}{2}\right)-1 \ge 0.95$$

于是 $\Phi\left(\dfrac{\sqrt{n}}{2}\right) \ge 0.975$，查表得 $\dfrac{\sqrt{n}}{2} \ge 1.96$，即 $n \ge (2\times 1.96)^2 = 15.3664$，所以 $n$ 的最小值 16。

12. 设总体 $X$ 服从 $(a,2)$ 上的均匀分布($a<2$ 并已知)，设 $X_1,X_2,X_3,X_4$ 是取自这个总体的一个样本，$\overline{X}$ 为样本均值。求：（1）$Y=3X_2$ 的概率密度;（2）相关系数 $\rho_{X_1X_3}$；（3）$\operatorname{cov}(X_2,\overline{X})$。

解：（1）$X_2$ 和 X 同分布，故 $f_{x_2}(x)=\begin{cases}1/(2-a) & a<x<2\\ 0 & \text{其他}\end{cases}$

因为 $y=g(x)=3x_2$，反函数 $x_2=h(y)=\dfrac{y}{3}$，$h'(y)=\dfrac{1}{3}$，$g'(x)=3$ 恒大于 0．由公式可知 $Y$ 的概率密度为 $f_Y(y)=f_X[h(y)]|h'(y)|=f_X\left(\dfrac{y}{3}\right)\dfrac{1}{|3|}=\dfrac{1}{3(2-a)}$, 3a<y<6

同时，$f_Y(y)=0$，y⩽3a 或 y⩾6

（2）由于独立性，则 $\rho_{X_1X_3}=0$

（3）

$$\begin{aligned}&\operatorname{cov}(X_2,\overline{X})=\operatorname{cov}(X_2,\frac{1}{4}\sum_{i=1}^{4}X_i)\\&=\frac{1}{4}\operatorname{cov}(X_2,X_1)+\frac{1}{4}\operatorname{cov}(X_2,X_2)+\frac{1}{4}\operatorname{cov}(X_2,X_3)+\frac{1}{4}\operatorname{cov}(X_2,X_4)\\&=\frac{1}{4}DX_2=\frac{(2-a)^2}{48}\end{aligned}$$

# 第七章 点估计：

13. 设总体 $X$ 的密度函数为 $f(x)=\begin{cases}\lambda a x^{a-1}e^{-\lambda x^a}, & x>0,\\ 0, & x\le 0.\end{cases}(\lambda>0,a>0)$，据来自总体 $X$ 的样本（$X_1$，$X_2$，$\cdots$，$X_n$），试求未知参数 $\lambda$ 的最大似然估计量。

解：最大似然函数为 $f(x_1,\cdots,x_n,\lambda)=\lambda^n a^n e^{-\lambda\sum_{i=1}^{n}x_i^a}\prod_{i=1}^{n}x_i^{a-1}$，取对数得：

$$\ln f=n\ln\lambda-\lambda\sum_{i=1}^{n}x_i^a-\ln\prod_{i=1}^{n}x_i^{a-1}$$

$$\frac{\partial\ln f}{\partial\lambda}=\frac{n}{\lambda}-\sum_{i=1}^{n}x_i^a=0。$$

所以 $\lambda = \dfrac{n}{\sum_{i=1}^{n} X_i^a}$。

14. 设某种电子元件的寿命 $T$ 服从参数 λ 的指数分布，今测得 10 个元件的失效时间为

1050，1100，1080，1200，1300，1250，1340，1060，1150，1150，

试求未知参数 λ 的最大似然估计值。

解：指数分布的密度函数为： $f(x,\lambda) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$,

作似然函数：$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$，取对数得：$\ln L(\lambda) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$，

令 $\dfrac{d \ln L(\lambda)}{d\lambda} = \dfrac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$，λ 的极大似然估计值为 $\hat{\lambda} = \dfrac{n}{\sum_{i=1}^{n} X_i} = \dfrac{1}{\overline{X}}$，

由样本的观测值得：

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10}(1050+1100+1080+1200+1300+1250+1340+1060+1150+1150)$$
$$=1168$$

所以，λ 的最大似然估计值为 $\hat{\lambda} = \dfrac{1}{1168} \approx 0.00086$。

15. 总体 $X$ 的密度函数 $f(x) = \begin{cases} \theta c^{\theta} x^{-\theta-1} & x > c \\ 0 & x \le c \end{cases}$ （$c > 0$ 为已知数），其中 $\theta > 1$ 为未知参数，$X_1, \cdots, X_n$ 为其样本，求 $\theta$ 的矩估计量。

解：$EX = \int_c^{+\infty} x f(x) \mathrm{d}x = \int_c^{+\infty} \theta c^{\theta} x^{-\theta} \mathrm{d}x = \theta c^{\theta} \cdot \dfrac{1}{-\theta+1} x^{-\theta+1} \Big|_c^{+\infty} = \dfrac{\theta c}{\theta - 1}$,

故矩估计量 $\hat{\theta} = \dfrac{\overline{X}}{\overline{X} - c}$。

16. 设总体 $X$ 的概率分布为

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

其中 $\theta$ $(0 < \theta < \dfrac{1}{2})$ 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 $\theta$ 的矩估计值和最大似然估计值。

解：$E(X) = 0 \times \theta^2 + 1 \times 2\theta(1-\theta) + 2 \times \theta^2 + 3 \times (1-2\theta) = 3 - 4\theta$，则

$\bar{x} = \dfrac{1}{8} \times ( 3 + 1 + 3 + 0 + 3 + 1 + 2 + 3,$

令 $E(X)=\bar{x}$，即 $3-4\theta=2$，解得 $\theta$ 的矩估计值为 $\theta=\dfrac{1}{4}$；

似然函数为：$L(\theta)=4\theta^6(1-\theta)^2(1-2\theta)^4$，得：

$$\ln L(\theta)=\ln 4+6\ln\theta+2\ln(1-\theta)+4\ln(1-2\theta),$$

即：$\dfrac{d\ln L(\theta)}{d\theta}=\dfrac{6}{\theta}-\dfrac{2}{1-\theta}-\dfrac{8}{1-2\theta}=\dfrac{6-28\theta+24\theta^2}{\theta(1-\theta)(1-2\theta)}$，

令 $\dfrac{d\ln L(\theta)}{d\theta}=0$，解得 $\theta=\dfrac{7+\sqrt{13}}{12}$ 或 $\theta=\dfrac{7-\sqrt{13}}{12}$，

因为 $0<\theta<\dfrac{1}{2}$，所以 $\theta$ 的最大似然估计值为：$\theta=\dfrac{7-\sqrt{13}}{12}$。

17. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的一个样本，$X$ 的概率密度为：

$$f(x)=\begin{cases}\dfrac{1}{2\theta} & -\theta<x<\theta \\ 0 & \text{其它}\end{cases},$$

其中 $\theta\,(\theta>0)$ 为未知参数。试求参数 $\theta$ 的矩估计量。

解：$X\sim U\ -\theta,\ \theta\ \Rightarrow EX=0$，$E\ X^2\ =\int_{-\theta}^{\theta}x^2\cdot\dfrac{1}{2\theta}dx=\dfrac{\theta^2}{3}\triangleq\dfrac{1}{n}\sum_{i=1}^{n}X_i^2$，

可以得到 $\theta$ 的矩估计量 $\hat{\theta}=\sqrt{\dfrac{3}{n}\sum_{i=1}^{n}X_i^2}$；

另解：$DX=\dfrac{2\theta\ ^2}{12}=\dfrac{\theta^2}{3}\triangleq S_n^2=\dfrac{1}{n}\sum_{i=1}^{n}\ X_i-\bar{X}\ ^2$，

整理可得，$\theta$ 的矩估计量 $\hat{\theta}=\sqrt{3}S_n=\sqrt{\dfrac{3}{n}\sum_{i=1}^{n}\ X_i-\bar{X}\ ^2}$。

18. 设 $X_1,X_2,\cdots,X_n$ 是来自总体 $X$ 的一个样本，$X$ 的概率密度为：

$$f(x,a)=\begin{cases}ax^{-2} & 0<a<x<\infty \\ 0 & \text{其它}\end{cases},$$

其中 $a$ 是未知参数，若得样本值为：0.1，0.2，0.9，0.8，0.7，0.7，试求参数 $a$ 的极大似然估计值。

解：似然函数为 $L\ a\ =\prod_{i=1}^{n}f\ x_i;a\ =\begin{cases}a^n\Big/\prod_{i=1}^{n}x_i^2, & \forall 0<a<x_i \\ 0, & \text{其他}\end{cases}$，$\forall 0<a<x_i$ 可以等价表示为 $0<a\le x_1\le x_2\le\ldots\le x_n$，故 $a$ 的极大似然估计量 $\hat{a}=X_1$，相应的估计值为 $\hat{a}=0.1$。

C

## 第五章 中心极限定理：

1. 抽样检查产品质量时，如果发现次品多于 10 个，则拒绝接受这批产品，设某批产品的次品率为 10%，问至少应抽取多少个产品检查才能保证拒绝接收该产品的概率达到 0.9？

解：设 $n$ 为至少应抽取的产品数，$X$ 为其中的次品数：

$$X_k=\begin{cases}1 & 第k次检查时为次品,\\ 0 & 第k次抽查时为正品,\end{cases}$$

则 $X=\sum_{k=1}^{n}X_k$，$E(X_k)=0.1$，$D(X_k)=0.1(1-0.1)=0.09$。

由棣莫弗-拉普拉斯中心极限定理，有

$$P\{X>10\}=P\{\frac{X-n\times 0.1}{\sqrt{n\times 0.1\times 0.9}}>\frac{10-n\times 0.1}{\sqrt{n\times 0.1\times 0.9}}\}\approx 1-\Phi(\frac{10-0.1n}{0.3\sqrt{n}})。$$

由题意 $1-\Phi(\frac{10-0.1n}{0.3\sqrt{n}})=0.9 \ \Rightarrow\ \Phi(\frac{10-0.1n}{0.3\sqrt{n}})=0.1$，查表得

$\frac{10-0.1n}{0.3\sqrt{n}}=-1.282 \ \Rightarrow\ n=147$。

## 第六章 样本及抽样分布：

2. 在总体 $N(12,4)$ 中随机抽取一容量为 3 的样本 $X_1,X_2,X_3$。试求：（1）$P\left\{\left|\overline{X}-12\right|<1\right\}$；（2）$P\left\{\max\left(X_1,X_2,X_3\right)>15\right\}$。

解：（1）由于 $\overline{X}\sim N\left(12,\frac{4}{3}\right)$，所以

$$P\left\{\left|\overline{X}-12\right|<1\right\}=P\left\{\frac{\left|\overline{X}-12\right|}{\sqrt{4/3}}<\frac{1}{\sqrt{4/3}}\right\}=\Phi\left(\frac{1}{\sqrt{4/3}}\right)-\Phi\left(\frac{-1}{\sqrt{4/3}}\right)=2\Phi\left(\frac{\sqrt{3}}{2}\right)-1=0.6156；$$

（2）$X_{(3)}=\max\left(X_1,X_2,X_3\right)\sim F_{X_{(3)}}\left(x\right)=\left[F_X\left(x\right)\right]^3=\left[\Phi\left(\frac{x-12}{2}\right)\right]^3$，所以

$$P\left\{\max\left(X_1,X_2,X_3\right)>15\right\}=1-F_{X_{(3)}}\left(15\right)=1-\left[\Phi\left(\frac{15-12}{2}\right)\right]^3=1-\left[\Phi\left(1.5\right)\right]^3$$

$$=1-0.9332^3=0.1873。$$

3. 在总体 $X\sim N(12,2^2)$ 中随机抽取一个容量为 5 的样本 $X_1,X_2,\cdots,X_5$，其顺序统计量为 $X_{(1)},X_{(2)},\cdots,X_{(5)}$，试求（1）$P\left\{X_{(5)}<15\right\}$；（2）$P\left\{X_{(1)}<10\right\}$。

解：因为 $X_1,X_2,\cdots,X_5$ 是来自总体 $X$ 的一个样本，而 $X\sim N(12,2^2)$，故 $X_i\sim N(12,2\ )^2$（$1\le i\le 5$）， （1）$P\left\{X_{(5)}<15\right\}=P\left\{\max\limits_{1\le i\le 5}\left\{X_i\right\}<15\right\}=\left[P\left\{X<15\right\}\right]^5=\left[F(15)\right]^5$

$$=\left[\Phi(\frac{15-12}{2})\right]^5=\left[\Phi(1.5)\right]^5=0.9332^5=0.7077；$$

（2）$P\left\{X_{(1)}<10\right\}=P\left\{\min\limits_{1\le i\le 5}\left\{X_i\right\}<10\right\}=1-\left[1-P\left\{X<10\right\}\right]^5=1-\left[1-F(10)\right]^5$

$$=1-\left[1-\Phi(\frac{10-12}{2})\right]^5=1-\left[\Phi(1)\right]^5=1-0.8413^5$$ 。

# 第七章 点估计：

4. 设某种元件的使用寿命 $X$ 的概率密度为 $f(x;\theta)=\begin{cases}2e^{-2(x-\theta)} & x>\theta\\ 0 & x\le\theta\end{cases}$，其中 $\theta>0$ 为未知参数，又设 $x_1,x_2,\cdots,x_n$ 是 $X$ 的一组样本观测值，试求（1）参数 $\theta$ 的最大似然估计量 $\hat\theta$；（2）求总体 $X$ 的分布函数 $F(x)$；（3）求统计量 $\hat\theta$ 的分布函数 $F_{\hat\theta}(x)$；（4）如果用 $\hat\theta$ 作为 $\theta$ 的估计量，讨论它是否具有无偏性。

解：（1）由 $X$ 的概率密度函数，得似然函数

$$L(\theta)=\prod_{i=1}^n f(x_i;\theta)=\prod_{i=1}^n 2e^{-2(x_i-\theta)}=2^n e^{-2\sum\limits_{i=1}^n(x_i-\theta)}=2^n e^{-2\sum\limits_{i=1}^n x_i+2n\theta}\quad x_i>\theta\ (i=1,2,\cdots,n),$$

取对数得： $\ln L(\theta)=n\ln 2-2\sum\limits_{i=1}^n x_i+2n\theta \quad x_i>\theta\ (i=1,2,\cdots,n)$，

再对 $\theta$ 求导得：$\dfrac{d\ln L(\theta)}{d\theta}=2n>0 \quad x_i>\theta(i=1,2,\cdots,n)$，

即 $L(\theta)$ 是单调增加的，虽然 $\theta$ 越大则 $L(\theta)$ 越大，但 $\theta$ 必须满足条件

$$\theta<x_i\ (i=1,2,\cdots,n)$$

所以当取 $\theta$ 为 $x_1,x_2,\cdots,x_n$ 中最小值 $x_{(1)}$ 时，$L(\theta)$ 取得满足条件的最大值，所以 $\theta$ 的最大似然估计值为 $\hat\theta=x_{(1)}=\min\{x_1,x_2,\cdots,x_n\}$。

（2）总体 $X$ 的分布函数 $F(x)=\int_{-\infty}^x f(t;\theta)dt$

当 $x\le\theta$ 时，$F(x)=\int_{-\infty}^x f(t;\theta)dt=0$；

当 $x>\theta$ 时，$F(x)=\int_{-\infty}^x f(t;\theta)dt=\int_\theta^x 2e^{-2(t-\theta)}dt=1-e^{-2(t-\theta)}$；

即得 $F(x)=\begin{cases}1-e^{-2(x-\theta)} & x>\theta\\ 0 & x\le\theta\end{cases}$；

（3）$\hat{\theta}=\min\{X_1,X_2,\cdots,X_n\}$ 的分布函数 $F_{\hat{\theta}}(x)$ 为：

$$F_{\hat{\theta}}(x)=P\{\hat{\theta}\le x\}=P\{\min\{X_1,X_2,\cdots,X_n\}\le x\}=1-P\{\min\{X_1,X_2,\cdots,X_n\}>x\}$$

$$=1-P\{X_1>x\}P\{X_2>x\}\cdots P\{X_n>x\}=1-\prod_{i=1}^{n}P\{X_i>x\}$$

$$=1-\prod_{i=1}^{n}[1-P\{X_i\le x\}]=1-\prod_{i=1}^{n}[1-F(x)]=1-[1-F(x)]^n=\begin{cases}1-e^{-2n(x-\theta)} & x>\theta\\ 0 & x\le\theta\end{cases};$$

（4）因为 $\hat{\theta}=\min\{X_1,X_2,\cdots,X_n\}$ 的概率密度函数为

$$f_{\hat{\theta}}(x)=F'_{\hat{\theta}}(x)=\begin{cases}2ne^{-2n(x-\theta)} & x>\theta\\ 0 & x\le\theta\end{cases}$$

而 $E(\hat{\theta})=\int_{\theta}^{+\infty}x\bullet 2ne^{-2n(x-\theta)}dx=-xe^{-2n(x-\theta)}\big|_{\theta}^{+\infty}+\int_{\theta}^{+\infty}e^{-2n(x-\theta)}dx=\theta+\dfrac{1}{2n}\neq\theta$，

所以 $\hat{\theta}=\min\{X_1,X_2,\cdots,X_n\}$ 不是 $\theta$ 的无偏估计。

5. 设总体 $X$ 服从伽玛分布，其概率密度函数为：

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{\beta^{-\alpha}}{\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}} & x>0\\ 0 & x\le 0\end{cases}\quad（\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}e^{-x}dx），$$

其中 $\alpha$ 是已知正实数，参数 $\beta$ $(\beta>0)$ 未知，若 $X_1,X_2,\cdots,X_n$ 是来自该总体的一个容量为 $n$ 的样本，试求参数 $\beta$ 的极大似然估计量，且问其是否无偏估计？

解：若总体 $X$ 的一个样本值为 $x_1,x_2,\cdots,x_n$，则本题总体的似然函数为：

$$L(\beta)=\prod_{i=1}^{n}f\left(x_i;\beta\right)=\prod_{i=1}^{n}\frac{\beta^{-\alpha}}{\Gamma(\alpha)}x_i^{\alpha-1}e^{-\frac{x_i}{\beta}}=\left(\frac{\beta^{-\alpha}}{\Gamma(\alpha)}\right)^n\left(\prod_{i=1}^{n}x_i\right)^{\alpha-1}e^{-\frac{1}{\beta}\sum_{i=1}^{n}x_i},$$

取对数得

$$\ln L(\beta)=-n\alpha\ln\beta-n\ln\Gamma(\alpha)+\left(\alpha-1\right)\ln\left(\prod_{i=1}^{n}x_i\right)-\frac{1}{\beta}\sum_{i=1}^{n}x_i,$$

再对 $\beta$ 求导数，并令其为 0：

$$\frac{d\ln L(\beta)}{d\beta}=-\frac{n\alpha}{\beta}+\frac{1}{\beta^2}\sum_{i=1}^{n}x_i=0,$$

解之得参数 $\beta$ 的极大似然估计值为：$\hat{\beta}=\dfrac{1}{n\alpha}\sum_{i=1}^{n}x_i=\dfrac{\overline{x}}{\alpha}$，

于是参数 $p$ 的极大似然估计量为：$\hat{\beta}=\dfrac{\bar{X}}{\alpha}$，

而 $X\sim\Gamma(\alpha,1/\beta)$，故期望 $E(X)=\alpha\beta$， $E(\bar{X})=\alpha\beta$，所以

$E\left(\hat{\beta}\right)=E\left(\dfrac{\bar{X}}{\alpha}\right)=\dfrac{1}{\alpha}E\left(\bar{X}\right)=\dfrac{1}{\alpha}\alpha\beta=\beta$，即 $\hat{\beta}=\dfrac{\bar{X}}{\alpha}$ 是参数 $\beta$ 的无偏估计。

6. 设自总体概率密度函数为 $f(x)=\begin{cases}\dfrac{1}{k\theta} & \theta<x<(k+1)\theta, \\ 0 & \text{其它}\end{cases}$，其中 $k$ 是一个已知的正常数，$\theta>0$ 是一个待估计的参数，样本 $x_1,x_2\cdots x_n$ 是来自总体的一个随机样本，（1）试求 $\theta$ 的极大似然估计 $\hat{\theta}_{极}$ 和矩估计 $\hat{\theta}_{矩}$；（2）$\hat{\theta}_{极}$ 是不是 $\theta$ 的无偏估计。

解：（1）$l(x_1,x_2\cdots x_n,\theta)=(\dfrac{1}{k\theta})^n$，

$Ln[l(x_1,x_2\cdots x_n,\theta)]=-n\ln k-n\ln\theta$，$\dfrac{\partial LnL}{\partial\theta}=-\dfrac{n}{\theta}<0$；

$\theta\le X_{(1)}\le X_{(2)}\cdots\le X_{(n)}\le(k+1)\theta\to\dfrac{X_{(n)}}{k+1}\le\theta\le X_{(1)}$

$\hat{\theta}_{极}=\dfrac{X_{(n)}}{k+1}$， $EX=\dfrac{k+2}{2}\theta\doteq\bar{X}$，$\hat{\theta}_{矩}=\dfrac{2}{k+2}\bar{X}$；

（2）$F_n(z)=P(X_{(n)}\le z)=(F_X(z))^n=\begin{cases}0 & z<\theta \\ (\dfrac{z-\theta}{k\theta})^n & \theta\le z<(k+1)\theta, \\ 1 & z>(k+1)\theta\end{cases}$

$$f_n(z)=\begin{cases}\dfrac{n(z-\theta)^{n-1}}{(k\theta)^n} & \theta\le z<(k+1)\theta \\ 0 & \text{其他}\end{cases}$$

$$E\hat{\theta}_{极}=\frac{1}{k+1}\int_{\theta}^{(k+1)\theta}z\frac{n(z-\theta)^{n-1}}{(k\theta)^n}dz=\frac{nk\theta}{n+1}+\frac{\theta}{k+1}$$

故 $\hat{\theta}_{极}$ 不是 $\theta$ 的无偏估计。

7. 设总体 $X$ 的分布函数为 $F(x)=\begin{cases}1-e^{-\frac{x^2}{\theta}} & x>0 \\ 0 & x\le 0\end{cases}$， $\theta>0$ 为未知参数，$X_1,X_2,\cdots,X_n$ 为来自总体 $X$ 的简单随机样本。（1）求 $E(X)$ 及 $E(X^2)$；（2）求 $\theta$ 的最大似然估计量 $\hat{\theta}$。

解：（1） $X$ 的概率密度函数为 $f(x)=\begin{cases}\dfrac{2x}{\theta}e^{-\frac{x^2}{\theta}} & x>0\\ 0 & x\le 0\end{cases}$ ，

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_0^{+\infty}x\cdot\frac{2x}{\theta}e^{-\frac{x^2}{\theta}}dx=-xe^{-\frac{x^2}{\theta}}\Big|_0^{+\infty}+\int_0^{+\infty}e^{-\frac{x^2}{\theta}}dx$$，作换元

$\dfrac{x}{\sqrt{\theta}}=\dfrac{t}{\sqrt{2}}$，得 $EX=\displaystyle\int_0^{+\infty}e^{-\frac{x^2}{\theta}}dx=\int_0^{+\infty}e^{-\frac{t^2}{2}}d\frac{\sqrt{\theta}}{\sqrt{2}}t=\sqrt{\theta\pi}\int_0^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt=\frac{\sqrt{\theta\pi}}{2}$

$$E(X^2)=\int_{-\infty}^{+\infty}x^2f(x)dx=\int_0^{+\infty}x^2\cdot\frac{2x}{\theta}e^{-\frac{x^2}{\theta}}dx=-x^2e^{-\frac{x^2}{\theta}}\Big|_0^{+\infty}+\int_0^{+\infty}2xe^{-\frac{x^2}{\theta}}dx$$

$$=\theta\int_0^{+\infty}\frac{2x}{\theta}e^{-\frac{x^2}{\theta}}dx=\theta\int_{-\infty}^{+\infty}f(x)dx=\theta$$

（2）1）似然函数 $L(\theta)=\displaystyle\prod_{i=1}^{n}f(x_i;\theta)=\prod_{i=1}^{n}\frac{2x_i}{\theta}e^{-\frac{x_i^2}{\theta}}=\frac{2^n\prod_{i=1}^{n}x_i}{\theta^n}e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i^2}$ ；

2）取对数 $\ln L(\theta)=n\ln 2+\ln(\displaystyle\prod_{i=1}^{n}x_i)-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}x_i^2$ ；

3）求导数 $\dfrac{d\ln L(\theta)}{d\theta}=-\dfrac{n}{\theta}+\dfrac{1}{\theta^2}\displaystyle\sum_{i=1}^{n}x_i^2\overset{令}{=}0$ ；

4）解之得 $\hat{\theta}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}x_i^2=\overline{x^2}$ 为 $\theta$ 的最大似然估计值；

5）$\theta$ 的最大似然估计量 $\hat{\theta}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i^2=\overline{X^2}$ 。