概率论与数理统计

1
$$\begin{cases} P(A\cup B)=P(A)+P(B)-P(AB)=0.6 \\ P(A)+P(B)=0.8 \end{cases}$$

由上式解出$P(AB)=0.2$。所以：

$$\begin{aligned} P(B\overline{A})+P(\overline{B}A) &= P(B(S-A))+P(A(S-B))=P(B-AB)+P(A-B) \\ &= P(A)+P(B)-2P(AB)=0.8-2\times 0.2=0.2 \end{aligned}$$

2 因为$P(A|B)=P(B|A)=0.4$，所以有：

$$\frac{P(AB)}{P(B)}=\frac{P(AB)}{P(A)}=0.4$$

所以$P(A)=P(B)$，且$P(AB)=0.4P(B)$

又因为$P(\overline{A}|\overline{B})=0.7$，所以：

$$P(\overline{A}|\overline{B})=\frac{P(\overline{A}\,\overline{B})}{P(\overline{B})}=0.7$$

根据德摩根定律：$\overline{A}\cap\overline{B}=\overline{A\cup B}$，有：

$$P(\overline{A}|\overline{B})=\frac{P(\overline{A}\,\overline{B})}{P(\overline{B})}=\frac{P(\overline{A\cup B})}{1-P(B)}=\frac{1-P(A\cup B)}{1-P(B)}=0.7$$

所以，有$P(A\cup B)=0.3-0.7P(B)$，又因为：

$$P(A\cup B)=P(A)+P(B)-P(AB)=2P(B)-0.4P(B)=1.6P(B)$$

由上述两式，联立解得$P(A+B)=P(A\cup B)=8/15$

3 因为$P(B)=0.5$，所以$P(\overline{B})=P(B)$，

又因为$P(A|B)=P(A|\overline{B})$，因此：

$$\frac{P(AB)}{P(B)}=\frac{P(A\overline{B})}{P(\overline{B})}$$

也就是$P(AB)=P(A\overline{B})$，也就是

$$P(AB)=P(A\overline{B})=P(A(S-B))=P(A-AB)=P(A)-P(AB)$$

即$P(AB)=0.5P(A)=0.2$，即$P(A\overline{B})=0.2$。

4 根据题意，只有A发生的概率即为$P(A\overline{B})=0.4$，只有B发生的概率即为$P(\overline{A}B)=0.4$。又因为A、B相互独立，所以A、$\overline{B}$和$\overline{A}$、B也是相互独立的，所以：

$$P(A\overline{B})=P(A)P(\overline{B})=P(A)(1-P(B))=P(A)-P(A)P(B)=0.4$$

$$P(\overline{A}B)=P(\overline{A})P(B)=(1-P(A))P(B)=P(B)-P(A)P(B)=0.4$$

因此得出$P(A)=P(B)$，代入就有：

$P(A)-P^2(A)=0.4$，解出$P(A)=0.5$

5 设选取一等品的事件为A_1，二等品的事件为A_2，三等品的事件为A_3。因此从中选取结果不是三等奖的事件为A_1+A_2，选取一等品的事件为A_1，题意中要求即：

$$P(A_1|A_1+A_2)=\frac{P(A_1\cap(A_1\cup A_2))}{P(A_1\cup A_2)}=\frac{P((A_1\cap A_1)\cup(A_1\cap A_2))}{P(A_1\cup A_2)}$$
$$=\frac{P(A_1\cup(A_1\cap A_2))}{P(A_1\cup A_2)}$$

由于选取奖品的事件是互不相容的，所以三个事件没有交集，因此：（比如你抽取到一等品，与你抽取到二等品在集合上是没有关系的，也就是两者不可能同时发生）

$$\text{原式}=\frac{P(A_1\cup(A_1\cap A_2))}{P(A_1\cup A_2)}=\frac{P(A_1)+P(A_1A_2)-P(A_1A_2A_3)}{P(A_1)+P(A_2)-P(A_1A_2)}$$
$$=\frac{P(A_1)}{P(A_1)+P(A_2)}=\frac{0.6}{0.6+0.3}=2/3$$

6 X服从超几何分布；

$$P(X=2)=\frac{C_4^2C_6^0}{C_{10}^2}=2/15$$
$$E(X)=n\frac{M}{N}=2\times\frac{4}{10}=4/5$$

注：若X服从超几何分布，则$E(X)=n\frac{M}{N}$，其中$\frac{M}{N}$表示次品率，题中次品率为4/10。

7 设三重伯努利试验的事件为X，设事件A的概率为p，由题意可知$X\sim b(3,p)$。A至少发生一次的概率为：

$$P(X\geq 1)=1-P(X=0)=1-C_3^0p^3=19/27$$

解得$p=2/3$。

8 根据二项分布的数学期望公式和方差公式，有$E(X)=np=3\times 0.5$、$D(X)=np(1-p)=3\times 0.5\times 0.5=0.75$，且：

$$P(X\geq 1)=1-P(X=0)=1-C_3^0p^3=1-(\frac{1}{2})^3=\frac{7}{8}$$

9 设100重伯努利试验的事件为X，易知$X\sim b(100,p)$，其方差为$D(X)=np(1-p)=100p-100p^2$，其一阶导数和二阶导数为：

$$\frac{d}{dp}D(X)=100-200p$$
$$\frac{d^2}{dp^2}D(X)=-200<0$$

求得驻点为$p=0.5$，且由于二阶导数在该驻点处小于0，可以确定这个驻点就是方差的极大值点。因此当$p=0.5$时，方差取得极大值，且极大值为$D(X)_{max}=100\times 0.5\times 0.5=25$

10 引入：

考虑进行n次伯努利实验，若第n次成功了，那么前面的$n-1$次试验均失败，也就是$P(X=n)=p(1-p)^{1-n}$次。特别的，我们称X服从几何分布，即$X\sim G(p)$
。

根据题意，我们知道X服从几何分布，它的表达式为：

$$P(X=k)=\frac{3}{4}(\frac{1}{4})^{k-1}$$

当X取偶数的时候，设概率为p，则：

$$\begin{aligned}p&=P(X=2)+P(X=4)+P(X=6)+...+P(X=2k)\\&=\frac{3}{4}(\frac{1}{4})+\frac{3}{4}(\frac{1}{4})^3+\frac{3}{4}(\frac{1}{4})^5+...+\frac{3}{4}(\frac{1}{4})^{2k-1}\\&=\frac{3}{4}[(\frac{1}{4})+(\frac{1}{4})^3+(\frac{1}{4})^5+...+(\frac{1}{4})^{2k-1}]\end{aligned}$$

设等比数列：

$$a_k=(\frac{1}{4})^{2k-1}$$

可以知道：

$$\begin{aligned}S_k&=(\frac{1}{4})+(\frac{1}{4})^3+(\frac{1}{4})^5+...+(\frac{1}{4})^{2k-1}\\&=\frac{(1/4)(1-(1/4)^{2k})}{1-(1/4)^2}=\frac{4}{15}[1-(1/16)^k]\end{aligned}$$

所以

$$p=\frac{3}{4}\times\frac{4}{15}[1-(1/16)^k]=\frac{3}{15}[1-(1/16)^k]$$

> 对于高中生而言，在这里就可以得到满分了
>
> 不过，严格来讲：
>
> $$p=\lim_{k->\infty}\frac{3}{15}[1-(1/16)^k]=3/15$$
>
> 所以最终答案应该是$3/15$，不过只要能写出p的表达式就算正确。

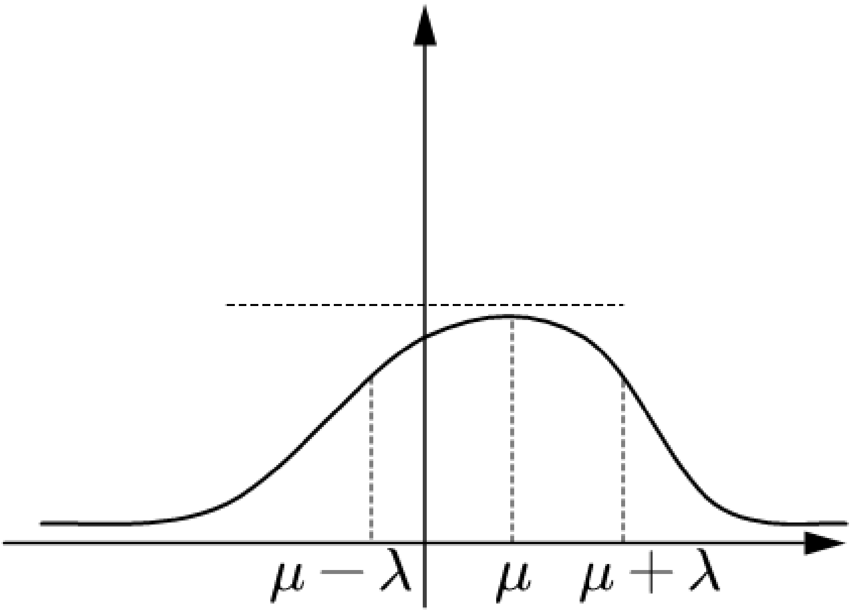
11 由于$X\sim N(1,4)$不是标准化的正态分布，所以将X标准化：

$$\frac{X-1}{2}\sim N(0,1)$$

因此：

$$P(-5<X<7)=P(\frac{-5-1}{2}<\frac{X-1}{2}<\frac{7-1}{2})=P(-1<\frac{X-1}{2}<1)=0.9974$$

12 根据正态分布的对称性：



设正态分布的概率密度函数为$f(x)$，所以我们知道：

$$P(X>\mu+\lambda)=\int_{\mu+\lambda}^{\infty}f(x)dx=p$$

由对称性可知：

$$P(X<\mu-\lambda)=\int_{-\infty}^{\mu-\lambda}f(x)dx=p$$

根据：

$$\int_{-\infty}^{\infty}f(x)dx=\int_{-\infty}^{\mu-\lambda}f(x)dx+\int_{\mu-\lambda}^{\mu+\lambda}f(x)dx+\int_{\mu+\lambda}^{\infty}f(x)dx=1$$

即：

$$\int_{\mu-\lambda}^{\mu+\lambda}f(x)dx=1-2p=P(\mu-\lambda<X<\mu+\lambda)$$

13 根据题意，可以知道：$X_i\sim b(100,0.5)$

$$\mu=E(X_i)=100\times0.5=50$$
$$\sigma^2=D(X_)=100\times0.5\times0.5=25$$

因此

$$\sum_{i=1}^{100}X_i\sim N(100\times50,100\times25)\sim N(5000,2500)$$

令$\Sigma_{i=1}^{100}X_i=Z$，则随机变量$Z\sim N(5000,2500)$

所以：

$$P(5000<\sum_{i=1}^{100}X_i<5050)=P(5000<Z<5050)$$

将Z标准化，也就是：

$$\frac{Z-5000}{50}\sim N(0,1)$$

也就是：

$$P(5000 < \sum_{i=1}^{100} X_i < 5050) = P(5000 < Z < 5050)$$
$$= P(\frac{5000-5000}{25} < \frac{Z-5000}{50} < \frac{5050-5000}{50})$$
$$= P(0 < \frac{Z-5000}{50} < 1) = 0.6826/2 = 0.3413$$

14 根据题目意思：

$$P(X > 3) = \int_3^{+\infty} 3e^{-3x}dx = e^{-9}$$

15 不管前面有多少人抽球，但每个人在抽球之后都放回了，所以第i个人抽球的时候，袋中仍然有a个白球和b个红球。因此：

$$p = \frac{C_a^1}{C_{a+b}^1} = \frac{a}{a+b}$$

16 在2000中，一共有$[2000/6] = 333$个数可以被6整除，一共有$[2000/8] = 250$个数可以被8整除。

因此，设事件A是取出来的数可以被6整除，事件B是取出来的数可以被8整除：

$$P(A) = \frac{333}{2000}$$
$$P(B) = \frac{250}{2000}$$

那么取到的整数既不能被6整除也不能被8整除的概率即为$P(\overline{A} \cap \overline{B})$，根据德摩根定律：

$$P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$$
$$= 1 - P(A) - P(B) + P(A \cap B)$$

而$A \cap B$表示的是取出的整数既可以被6整除，又可以被8整除。

根据最小公倍数和最大公约数的关系，我们首先求出6和8的最大公约数：

$$\gcd(6,8) = \gcd(8,6) = \gcd(6,2) = \gcd(2,0) = 2$$

所以6和8的最大公约数为2。

然后求出最小公倍数：

$$\text{lcm}(6,8) = \frac{6 \times 8}{\gcd(6,8)} = 48/2 = 24$$

所以6和8的最小公倍数为24.

因此，既然这个整数不能被6和8整除，也就是不能被24整除。在2000个数中，一共有$[2000/24] = 83$个数可以被24整除。所以：

$$P(A \cap B) = \frac{83}{2000}$$

所以代入，解出：

$$P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$$
$$= 1 - P(A) - P(B) + P(A \cap B) = 1 - \frac{333}{2000} - \frac{250}{2000} + \frac{83}{2000} = \frac{3}{4}$$

> $f(x,y) = \gcd(x,y)$是计算最大公约数的函数，它是一种离散型二元函数，属于数论领域，并且十分广泛用于计算机科学。

我们是利用辗转相除法这个算法来实现gcd的运算，具体算法是这样的：

将x和y进行位置互换，并且第二位置变成$x\%y$，也就是：

$f(x,y)=\gcd(x,y)=\gcd(y,x\%y)$

其中$x\%y$的意思是求出x/y的余数，比如3/2的余数就是1，也就是$3\%2=1$。

这个过程将继续下去，直到第二位置为0，也就是：

$f(x,y)=\gcd(x,y)=\gcd(y,x\%y)=\gcd(x\%y,y\%(x\%y))=\ldots=\gcd(g,0)=g$

此时g就是x和y的最大公约数。

最小公倍数的计算相对简单，直接利用公式：

$$\operatorname{lcm}(x,y)=\frac{xy}{\gcd(x,y)}$$

因为计算最小公约数的时候会利用递归，所以一般交给计算机计算。你可以尝试用C语言程序来编写一个计算最小公约数和最大公倍数的算法吗？（入门程序）

17 利用集合论方法：

首先找到事件的全集，也就是样本空间：

$S=\{HH,HT,TH,TT\}$

然后是事件A：

$A=\{HH,HT,TH\}$

事件B：

$B=\{HH,TT\}$

所以：

$A\cap B=\{HH\}$

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{1/4}{3/4}=1/3$$

也可以利用古典概型：

计算出集合之后，知道A集合有3个元素，B集合有2两个元素，AB集合有1个元素，因此：

$$P(B|A)=\frac{card(A\cap B)}{card(A)}=\frac{1}{3}$$

18 第一次取到的是一等品的概率：

$$P(A)=\frac{C_3^1}{C_4^1}=3/4$$

第二次取到的是一等品的概率：

$$P(B)=\frac{C_3^1}{C_4^1}\frac{C_2^1}{C_3^1}+\frac{C_1^1}{C_4^1}\frac{C_3^1}{C_3^1}=3/4$$

又因为AB事件的含义是第一次抽到一等品，第二次抽到一等品，所以：

$$P(AB)=\frac{C_3^1}{C_4^1}\frac{C_2^1}{C_3^1}=1/2$$

所以：

$$P(B|A)=\frac{P(AB)}{P(A)}=\frac{1/2}{3/4}=2/3$$

19 设A事件为电子设备厂中产品的次品率，B_1表示元件来自1厂，B_2表示元件来自2厂，B_3表示元件来自3厂。根据表格知道：

$$\begin{aligned}P(A|B_1)&=0.02 \quad P(B_1)=0.15\\P(A|B_2)&=0.01 \quad P(B_2)=0.80\\P(A|B_3)&=0.03 \quad P(B_3)=0.05\end{aligned}$$

（1）根据全概率公式，有：

$$\begin{aligned}P(A)&=P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+P(A|B_3)P(B_3)\\&=0.02\times 0.15+0.01\times 0.80+0.03\times 0.05=0.0125\end{aligned}$$

（2）根据贝叶斯公式，有：

$$P(B_1|A)=\frac{P(AB_1)}{P(A)}=\frac{P(A|B_1)P(B_1)}{P(A)}=\frac{0.02\times 0.15}{0.0125}=0.24$$

$$P(B_2|A)=\frac{P(AB_2)}{P(A)}=\frac{P(A|B_2)P(B_2)}{P(A)}=\frac{0.01\times 0.80}{0.0125}=0.64$$

$$P(B_3|A)=\frac{P(AB_3)}{P(A)}=\frac{P(A|B_3)P(B_3)}{P(A)}=\frac{0.03\times 0.05}{0.0125}=0.12$$

> 全概率公式：
>
> 若S为样本空间，且B_i为样本空间的划分，符合：
>
> $$\bigcup_{i=1}^{n}B_i=S$$
>
> $$B_iB_j=\emptyset$$
>
> 那么
>
> $$\begin{aligned}P(A)=P(AS)=P(A\cap(B_1\cup B_2\cup\ldots\cup B_n))&=P(AB_1)+P(AB_2)+\ldots+P((AB_n))\\&=P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+\ldots+P(A|B_n)P(B_n)\\&=\sum_{i=1}^{n}P(A|B_i)P(B_i)\end{aligned}$$
>
> 贝叶斯公式：（逆概公式）
>
> 同理：
>
> $$\begin{aligned}P(B_i|A)&=\frac{P(AB_i)}{P(A)}=\frac{P(A|B_i)P(B_i)}{P(A)}\\&=\frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n}P(A|B_j)P(B_j)}\end{aligned}$$

20 设机器调整好的事件为A，产品合格的事件为B。那么根据题意，有：

$P(B|A)=0.98$，$P(B|\overline{A})=0.55$。

早上启动机器的时候，有95%的概率机器调整良好，也就是在早上有

$P(A)=0.95$。

因此：根据全概率公式

$$P(B)=P(B|A)P(A)+P(B|\overline{A})P(\overline{A})=0.98\times 0.95+0.55\times 0.05=0.9585$$

那么根据贝叶斯公式，早上第一件产品是合格品时，机器调整良好的概率为：

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} = \frac{0.98 \times 0.95}{0.9585} = 0.9713$$

21 若采取三局两胜，那么甲胜利的情况有三种：

$\boxed{甲}\boxed{甲}$　　$\boxed{甲}\boxed{乙}\boxed{甲}$　　$\boxed{乙}\boxed{甲}\boxed{甲}$

p^2　　$p(1-p)\cdot p$　　$(1-p)\cdot p\cdot p$

也就是：

$P_1 = p^2 + p^2(1-p) + p^2(1-p) = p^2 + 2p^2(1-p)$

若采取五局三胜，那么甲胜利的情况：

3局 $\boxed{甲}\boxed{甲}\boxed{甲}$　　p^3

4局 $\boxed{\ }\boxed{\ }\boxed{\ }\boxed{甲}$　　$p\cdot C_3^2 p\cdot p(1-p)$

前3局甲胜2局　　$= C_3^2 p^2(1-p)$

5局 $\boxed{\ }\boxed{\ }\boxed{\ }\boxed{\ }\boxed{\ }\boxed{甲}$　　$p\cdot C_4^2 p\cdot p\cdot(1-p)(1-p)$

前4局甲胜2局　　$= C_4^2 p^3(1-p)^2$

因此：

$P_2 = p^3 + 3p^2(1-p) + 6p^3(1-p)^2$

做差，令函数

$f(p) = P_2 - P_1 = p^3 + p^2(1-p) - p^2 + 6p^3(1-p)^2 = 3p^2(p-1)^2(2p-1)$

定义域为$p \in [1/2, 1]$，所以：

当$p > 1/2$时，明显$f(p) > 0$，此时$P_2 > P_1$，这时五局三胜对甲更有利；

当$p = 1/2$时，明显$f(p) = 0$，此时$P_2 = P_1$，这时候无论三局两胜还是五局三胜效果都是一样的。

22 （1）A和B互不相容，也就是$P(AB) = 0$，那么：

$$P(A\overline{B}) = P(A(S-B)) = P(A) - P(AB) = P(A) = 1/2$$

（2）同理：

$$P(A\overline{B}) = P(A(S-B)) = P(A) - P(AB) = P(A) = 1/2 - 1/8 = 3/8$$

23 （1）设命中次数为X，易知X服从二项分布，即$X \sim b(400, 0.02)$

由于$n = 400$相对p而言大的很多，所以此时利用二项分布难以计算。此时的分布近似于泊松分布，$\lambda = np = 8$，也就是$X \sim \pi(8)$。

至少命中两次的概率为：

$$P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - \frac{8^0 e^{-8}}{0!} - \frac{8^1 e^{-8}}{1!} = 1 - 9e^{-8}$$

（2）可以知道次品率$p = 0.001$，次数$n = 1000$，此时X近似服从泊松分布，$\lambda = np = 1$，因此：

$$P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - \frac{1^0 e^{-1}}{0!} - \frac{1^1 e^{-1}}{1!} = 1 - 2e^{-1}$$

24 $d = 90^\circ C$，此时$X \sim N(90, 0.25)$，那么：

$$P(X < 89) = P(\frac{X-90}{0.5} < \frac{89-90}{0.5}) = P(\frac{X-90}{0.5} < -2) = \frac{1-0.9544}{2} = 0.0228$$

25 （1）

$$p = \frac{C_8^1}{C_{10}^1}\frac{C_7^1}{C_9^1} = 28/45$$

（2）

$$p = \frac{C_2^1}{C_{10}^1}\frac{C_1^1}{C_9^1} = 1/45$$

（3）

$$p = \frac{C_8^1}{C_{10}^1}\frac{C_2^1}{C_9^1} + \frac{C_2^1}{C_{10}^1}\frac{C_8^1}{C_9^1} = 16/45$$

（4）

$$p = \frac{C_8^1}{C_{10}^1}\frac{C_2^1}{C_9^1} + \frac{C_2^1}{C_{10}^1}\frac{C_1^1}{C_9^1} = 1/5$$

26 有两种情况：

一种是第一次取到白球，并放入到乙袋中；而第二种是第一次取到红球，并放入到乙袋中：

$$p = \frac{C_n^1}{C_{n+m}^1}\frac{C_{N+1}^1}{C_{N+M+1}^1} + \frac{C_m^1}{C_{n+m}^1}\frac{C_N^1}{C_{N+M+1}^1} = \frac{n(N+m+1)}{(n+m)(N+M+1)}$$

27 设三个人分别破译出密码的事件分别为A、B、C，那么三个人中至少有一个人能破译密码的概率即为：$P(A \cup B \cup C)$

而三个人中至少破译出密码的对立面就是三个人中没有人破译出密码，因此：

$$P(A \cup B \cup C) = 1 - P(\overline{A} \cap \overline{B} \cap \overline{C}) = 1 - P(\overline{A})P(\overline{B})P(\overline{C}) = 3/5 = 0.4$$

> 这个式子底层原理也是德摩根定律，具体来讲，这个式子应该是：
>
> $$P(A \cup B \cup C) = 1 - P(\overline{A \cup B \cup C}) = 1 - P(\overline{A} \cap \overline{B} \cap \overline{C})$$
>
> 所以，请牢记德摩根定律，这个定律在布尔代数和概率论中十分广泛，并且这个定律的现实意义解释也是十分清楚。
>
> 比如这道题，三个人中至少能有个人破译出密码，也就是1−三个人中没有人能破译出密码，也就是$1 - P(\overline{A} \cap \overline{B} \cap \overline{C})$

28 （1）根据题意，得到对数似然函数：

$$\ln L(\theta) = -\theta^2 + 2\theta - 1$$

所以：

$$\frac{d}{d\theta}\ln L(\theta) = -2\theta + 2 = 0$$

解得：

$$\theta = 1$$

为函数$\ln L(\theta)$的驻点

又因为：

$$\frac{d^2}{d\theta^2}\ln L(\theta) = -2 < 0$$

因此$\theta = 1$为函数$\ln L(\theta)$的极大值点

所以最大似然估计量为：

$$\hat{\theta} = 1$$

（2）不难得出似然函数为：

$$\begin{aligned} L(\theta) &= f(x_1;\theta)f(x_2;\theta)\ldots f(x_n;\theta) \\ &= \theta^n \prod_{i=1}^{n} x_i^{-(\theta+1)} \end{aligned}$$

然后求其对数似然函数：

$$\begin{aligned} \ln L(\theta) &= \ln \theta^n \prod_{i=1}^{n} x_i^{-(\theta+1)} \\ &= \ln \theta^n + \ln \prod_{i=1}^{n} x_i^{-(\theta+1)} \\ &= n\ln\theta + \ln\left(x_1^{-(\theta+1)} x_2^{-(\theta+1)} \ldots x_n^{-(\theta+1)}\right) \\ &= n\ln\theta - (\theta+1)(\ln x_1 + \ln x_2 + \ldots + \ln x_n) \\ &= n\ln\theta - (\theta+1)\sum_{i=1}^{n}\ln x_i \end{aligned}$$

其导数

$$\frac{d}{d\theta}\ln\theta = \frac{n}{\theta} - \sum_{i=1}^{n}\ln x_i = 0$$

解得：

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n}\ln x_i}$$