概率论与数理统计 第9次作业

Name: 宋昊原　　Student ID: 2022010755

November 18, 2023

1　概率知识总结

- 事件的概率
  - 事件、事件的运算
  - 概率的三种解释
    - 古典解释
    - 频率解释
    - 主观解释
  - 概率的公理化定义
  - 条件概率
  - 事件的独立性
  - Bayes公式
- 随机变量
  - 累积分布函数
  - 离散型随机变量和概率质量函数
    - Bernoulli分布
    - 二项分布
    - Poisson分布
  - 连续型随机变量和概率密度函数
    - 均匀分布
    - 正态分布
    - 指数分布
  - 随机变量的函数
- 随机向量
  - 联合累积分布函数
  - 离散分布和联合概率质量函数
    - 多项分布
  - 连续分布和联合概率密度函数
    - 二元正态分布
  - 边际分布
  - 条件分布和随机变量的独立性
  - 随机向量的函数
    - 直接求解函数的分布
    - 密度函数变换法

- 随机变量的数字特征
  - 期望
  - 分位数
  - 方差、标准差
  - 协方差、相关系数
  - 矩、偏度系数、峰度系数
  - 矩母函数
    * 矩母函数和矩的关系
    * 矩母函数确定分布
  - 条件期望和全期望公式
- 概率不等式与极限定理
  - 概率不等式
    * Markov不等式
    * Chebishev不等式
    * Chernoff不等式
  - 随机变量列的收敛性
    * 依概率收敛
    * 以概率1收敛
    * 依分布收敛
  - 极限定理
    * Khinchin弱大数定律
    * Kolmogrov强大数定律
    * 中心极限定理

主要问题在于期望方差的证明、随机变量列收敛的证明等证明题的思路.

2 抽样调查案例

高中数学课统计单元曾布置过一个作业，要求调查关于高考3+3选科方案的内容. 然而，我们抽样的方式采取发朋友圈的方便抽样的方式，由于课程关联、分班等原因，我们认识的同学中选择理科的比例要比年级里高得多，这样的抽样很可能不具备代表性.

3 简单随机抽样

3.1

只需证明X_i取得总体中任何一个选择x_j的概率确实是$\frac{1}{N}$即可.

$$P(X_i=x_j)=P(X_1\neq x_j\wedge...\wedge X_{i-1}\neq x_j)P(X_i=x_j|X_1\neq x_j\wedge...\wedge X_{i-1}\neq x_j)$$

$$=\frac{N-1}{N}\times\frac{N-2}{N-1}\times...\times\frac{N-i+1}{N-i+2}\times\frac{1}{N-i+1}=\frac{1}{N}$$

即所有X_i同分布，故有

$$E(X_i)=\sum_{j=1}^{N}\frac{1}{N}x_j=\mu$$

$$Var(X_i)=\sum_{j=1}^{N}\frac{1}{N}(x_j-\mu)^2=\sigma^2$$

3.2

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu$$

$$E(\bar{X}^2) = \frac{1}{n^2}(\sum_{i=1}^{n} E(X_i^2) + 2\sum_{1\leq i<j\leq n} E(X_iX_j))$$

其中

$$E(X_i^2) = Var(X_i) + E^2(X_i) = \mu^2 + \sigma^2$$

$$E(X_iX_j) = \frac{1}{N(N-1)}(\sum_{1\leq k,l\leq N\wedge k\neq l} x_kx_l)$$

$$= \frac{1}{N(N-1)}((\sum_{k=1}^{N} x_k)^2 - \sum_{k=1}^{N} x_k^2)$$

$$= \frac{1}{N(N-1)}(N^2\mu^2 - N(\mu^2+\sigma^2))$$

$$= \mu^2 - \frac{1}{N-1}\sigma^2$$

故

$$E(\bar{X}^2) = \frac{1}{n^2}(n(\mu^2+\sigma^2) + n(n-1)(\mu^2 - \frac{1}{N-1}\sigma^2))$$

$$= \mu^2 - \frac{1}{n}(\frac{N-n}{N-1})\sigma^2$$

故

$$Var(\bar{X}) = E(\bar{X}^2) - E^2(\bar{X}) = \frac{\sigma^2}{n}(\frac{N-n}{N-1})$$

4 二项总体的估计

4.1

总体的均值为kp，方差为$kp(1-p)$，矩估计为

$$kp \approx \bar{X}$$

$$kp(1-p) \approx \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

故

$$p \approx 1 - \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n\bar{X}}$$

$$k \approx \frac{\bar{X}}{1 - \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n\bar{X}}} = \frac{n\bar{X}^2}{n\bar{X} - \sum_{i=1}^{n}(X_i-\bar{X})^2}$$

4.2

根据第11题计算机实验中的讨论，上述估计在n或p较小时很不准确.

5 均匀总体的估计

5.1 矩估计

$$\mu = \frac{3}{2}\theta$$

故

$$\theta \approx \frac{2}{3}\bar{X}$$

5.2 极大似然估计

均匀分布的概率密度函数为

$$f_1(x;\theta) = \begin{cases} \frac{1}{\theta} & \theta \le x \le 2\theta \\ 0 & else \end{cases} \tag{1}$$

故似然函数

$$L(\theta) = \begin{cases} \frac{1}{\theta^n} & \theta \le X_1, ..., X_n \le 2\theta \\ 0 & else \end{cases} \tag{2}$$

要使$L(\theta)$取最大值，只需在满足$\theta \le X_1, ..., X_n \le 2\theta$的条件下取最小的$\theta$，即

$$\theta^* = \frac{\max\limits_i X_i}{2}$$

6 特殊分布的估计

6.1

显然有$f(x;a,\sigma) \ge 0, \forall x \in \mathbb{R}$

$$\int_{-\infty}^{+\infty} f(x;a,\sigma)dx = \frac{1}{\sigma^2}\int_{-\infty}^{+\infty} \frac{(x-a)^2}{\sqrt{2\pi}\sigma}\exp(-\frac{1}{2\sigma^2}(x-a)^2)dx$$
$$= \frac{1}{\sigma^2}Var(N(a,\sigma^2)) = 1$$

6.2 矩估计

总体均值

$$E(X) = \int_{-\infty}^{+\infty} \frac{x(x-a)^2}{\sqrt{2\pi}\sigma^3}\exp(-\frac{1}{2\sigma^2}(x-a)^2)dx$$
$$= \int_{-\infty}^{+\infty} \frac{(x-a)^3}{\sqrt{2\pi}\sigma^3}\exp(-\frac{1}{2\sigma^2}(x-a)^2)d(x-a) + a\int_{-\infty}^{+\infty} \frac{(x-a)^2}{\sqrt{2\pi}\sigma^3}\exp(-\frac{1}{2\sigma^2}(x-a)^2)dx$$
$$= 0 + a = a$$

总体方差

$$Var(X) = \int_{-\infty}^{+\infty} \frac{(x-a)^4}{\sqrt{2\pi}\sigma^3}\exp(-\frac{1}{2\sigma^2}(x-a)^2)dx$$
$$= \int_{-\infty}^{+\infty} \frac{x^4}{\sqrt{2\pi}\sigma^3}\exp(-\frac{1}{2\sigma^2}x^2)dx$$

令$\xi = \frac{x}{\sigma}$，有

$$Var(X) = \int_{-\infty}^{+\infty} \frac{\xi^4\sigma^2}{\sqrt{\pi}}\exp(-\frac{\xi^2}{2})d\xi = 3\sqrt{2}\sigma^2$$

故

$$a \approx \bar{X}$$
$$\sigma^2 \approx \frac{\sum\limits_{i=1}^{n}(X_i - \bar{X})^2}{3\sqrt{2}n}$$

6.3 极大似然估计

似然函数

$$L(a,\sigma^2)=\prod_{i=1}^{n}\frac{(X_i-a)^2}{\sqrt{2\pi}\sigma^3}\exp(-\frac{(X_i-a)^2}{2\sigma^2})$$

故

$$\ln L(a,\sigma^2)=\sum_{i=1}^{n}(2\ln|X_i-a|-\frac{(X_i-a)^2}{2\sigma^2}-\ln(\sqrt{2\pi}\sigma^3))$$

似然方程为

$$\frac{\partial \ln L(a,\sigma^2)}{\partial a}=\sum_{i=1}^{n}(-\frac{2}{X_i-a}+\frac{X_i-a}{\sigma^2})=0$$

$$\frac{\partial \ln L(a,\sigma^2)}{\partial \sigma^2}=\sum_{i=1}^{n}(\frac{(X_i-a)^2}{2\sigma^4}-\frac{3}{2\sigma^2})=0$$

即

$$\sum_{i=1}^{n}\frac{(X_i-a)^2}{\sigma^2}-2n=0$$

$$\sum_{i=1}^{n}\frac{(X_i-a)^2}{\sigma^2}-3n=0$$

这似乎是个矛盾方程.

7 Bernoulli分布估计

7.1 矩估计

$$p\approx\bar{X}$$

7.2 极大似然估计

$$L(p)=p^{X_1+...+X_n}(1-p)^{n-X_1-...-X_n}=p^{n\bar{X}}(1-p)^{n-n\bar{X}}$$

$$\ln L(p)=n\bar{X}\ln p+n(1-\bar{X})\ln(1-p)$$

令

$$\frac{\partial \ln p}{\partial p}=\frac{n\bar{X}}{p}-\frac{n(1-\bar{X})}{1-p}=0$$

$$(1-p^*)\bar{X}=p^*(1-\bar{X})$$

$$p^*=\bar{X}$$

这确实是最大值点.

8 多项分布极大似然估计

联合分布PMF为

$$L(\mathbf{p})=f(X_1,...,X_m;p_1,...,p_m)=\frac{n!}{X_1!...X_m!}p_1^{X_1}...p_m^{X_m}$$

$$\ln L(\mathbf{p})=\ln\frac{n!}{X_1!...X_m!}+X_1\ln p_1+...+X_m\ln p_m$$

需要找此函数在$p_1+...+p_m=1$条件下的条件极值，故考虑Lagrange函数:

$$\tilde{L}(\mathbf{p},\lambda)=\ln\frac{n!}{X_1!...X_m!}+X_1\ln p_1+...+X_m\ln p_m+\lambda(p_1+...+p_m-1)$$

故

$$\frac{\partial \ln L(\mathbf{p})}{\partial(\mathbf{p},\lambda)} = \begin{bmatrix} \frac{X_1}{p_1}+\lambda & ... & \frac{X_m}{p_m}+\lambda & p_1+...+p_m-1 \end{bmatrix}$$

令上述梯度为$\mathbf{0}$，则有

$$\frac{X_1}{p_1} = ... = \frac{X_m}{p_m} = -\lambda$$
$$p_1+...+p_m = 1$$

故

$$p_i = \frac{X_i}{X_1+...+X_m}$$

这是符合预期的估计.

9 给定分布的估计

9.1 矩估计

$$E(X) = \theta^2 + 4\theta(1-\theta) + 3(1-\theta)^2 = 3-2\theta$$

此时

$$\bar{X} = \frac{4}{3}$$

故

$$\hat{\theta} = \frac{5}{6}$$

9.2 极大似然估计

$$L(\theta) = 2\theta^5(1-\theta) = 2\theta^5 - \theta^6$$

令

$$\frac{dL(\theta)}{d\theta} = 5\theta^4 - 6\theta^5 = 0$$

有

$$\theta^* = \frac{5}{6}$$

这确实是最大值点.

10 某种幂分布的估计

10.1 矩估计

$$E(X) = \int_0^1 \theta x^\theta dx = \frac{\theta}{\theta+1}$$

故

$$\frac{\hat{\theta}}{\hat{\theta}+1} = \bar{X}$$
$$\hat{\theta} = \frac{\bar{X}}{1-\bar{X}}$$

10.2 极大似然估计

$$L(\theta) = \theta^n (\prod_{i=1}^{n} X_i)^{\theta-1}$$

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

令

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

解得

$$\theta^* = -\frac{n}{\sum\limits_{i=1}^{n} \ln X_i}$$

由于

$$L(0) = \lim_{\theta \to +\infty} L(\theta) = 0$$

知，这确实是最大值点.

11 计算机实验：矩估计

11.1 数据展示

```
[1] "参数：k=10,p=0.01，模拟次数10"
[1] "第1次模拟：p=NaN,k=NaN"
[1] "第2次模拟：p=NaN,k=NaN"
[1] "第3次模拟：p=NaN,k=NaN"
[1] "第4次模拟：p=0.0000000000,k=Inf"
[1] "第5次模拟：p=0.0000000000,k=Inf"
[1] "第6次模拟：p=0.2222222222,k=1.3500000000"
[1] "第7次模拟：p=NaN,k=NaN"
[1] "第8次模拟：p=0.0000000000,k=Inf"
[1] "第9次模拟：p=NaN,k=NaN"
[1] "第10次模拟：p=-0.2222222222,k=-1.8000000000"
```

```
[1] "参数：k=10,p=0.01，模拟次数1000"
[1] "第1次模拟：p=-0.0411893375,k=-2.6220378038"
[1] "第2次模拟：p=-0.0744955482,k=-1.2752439887"
[1] "第3次模拟：p=-0.0041152263,k=-26.2440000000"
[1] "第4次模拟：p=0.0049175389,k=20.9454367589"
[1] "第5次模拟：p=-0.0508842176,k=-1.8866360656"
[1] "第6次模拟：p=0.0116783450,k=8.2203428571"
[1] "第7次模拟：p=0.0396110396,k=2.8274945848"
[1] "第8次模拟：p=-0.0018895300,k=-47.1016607143"
[1] "第9次模拟：p=0.0172091283,k=5.7527608696"
[1] "第10次模拟：p=0.0604564961,k=1.6706227869"
```

```
[1] "参数：k=10,p=0.50，模拟次数10"
[1] "第1次模拟：p=0.3333333333,k=15.0000000000"
[1] "第2次模拟：p=0.6356589147,k=6.7646341463"
[1] "第3次模拟：p=0.7407407407,k=6.0750000000"
[1] "第4次模拟：p=0.6521739130,k=7.0533333333"
[1] "第5次模拟：p=0.2510288066,k=21.5114754098"
[1] "第6次模拟：p=0.3333333333,k=13.2000000000"
[1] "第7次模拟：p=0.5248226950,k=8.9554054054"
[1] "第8次模拟：p=0.2411347518,k=19.4911764706"
[1] "第9次模拟：p=0.0725623583,k=67.5281250000"
[1] "第10次模拟：p=0.6666666667,k=7.8000000000"
```

```
[1] "参数：k=10,p=0.50，模拟次数1000"
[1] "第1次模拟：p=0.5003635570,k=10.0506920004"
[1] "第2次模拟：p=0.5366186638,k=9.4033255642"
[1] "第3次模拟：p=0.5352055031,k=9.6056560893"
[1] "第4次模拟：p=0.4558843980,k=10.8316933460"
[1] "第5次模拟：p=0.5295505947,k=9.3324416012"
[1] "第6次模拟：p=0.4909248872,k=10.1481919748"
[1] "第7次模拟：p=0.4923854875,k=10.0693463260"
[1] "第8次模拟：p=0.5144872163,k=9.6931465776"
[1] "第9次模拟：p=0.4962459941,k=10.0010882883"
[1] "第10次模拟：p=0.5009835335,k=9.8945367839"
```

11.2 问题

在p很小或n很小时，很容易得到估计不准的情况，p小的时候参数可能会被计算成负值，n小的时候虽然不会出现负值但估计也很不准确，而n, p都很小的时候甚至会产生数学错误（发生0次时，均值和方差均为0，产生NaN，而发生1次时，p的估计值为0，k产生Inf）
只有p, n都足够大的情况下，矩估计才能得到好的估计结果