概率论与数理统计 (Fall 2024)
习题课讲义

2024 年 11 月 18 日

聂嵘琢, 祝国隽

B01GB006Y-02/Note9

---

1. 记 $X_1, X_2, \cdots, X_{2n}$ iid $\sim N(0,1)$, 求

$$Y = \frac{1}{2}\sum_{i=1}^{2n} X_i^2 + \sum_{i=1}^{n} X_{2i-1}X_{2i}$$

的分布.

Ans: 由于

$$\begin{aligned} Y &= \frac{1}{2}\sum_{i=1}^{2n} X_i^2 + \sum_{i=1}^{n} X_{2i-1}X_{2i} \\ &= \frac{1}{2}\left(X_1^2 + X_2^2 + \cdots + X_{2n}^2\right) + X_1X_2 + X_3X_4 + \cdots + X_{2n-1}X_{2n} \\ &= \frac{1}{2}\left(X_1 + X_2\right)^2 + \frac{1}{2}\left(X_3 + X_4\right)^2 + \cdots + \frac{1}{2}\left(X_{2n-1} + X_{2n}\right)^2 \\ &= \sum_{i=1}^{n}\left(\frac{X_{2i-1} + X_{2i}}{\sqrt{2}}\right)^2 \end{aligned}$$

且 $\dfrac{X_{2i-1} + X_{2i}}{\sqrt{2}} \sim N(0,1)$, 则 $Y \sim \chi_n^2$.

有些同学是这么做的, 以下是一组对比:

习题10. 令 $X = (X_1, X_2, \ldots, X_{2n})^T$，$X \sim N(0_{2n}, I_{2n})$.

令 $Z = \Delta^{\frac{1}{2}} X = (Z_1, Z_2, \ldots, Z_n)^T$，其中 $\Delta = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & \ddots & & \vdots \\ 0 & 0 & \ddots & \ddots & \frac{1}{2} \\ \vdots & \vdots & & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

则 $Z \sim N(0_{2n}, \Delta)$，且 $Y = \frac{1}{2}\sum_{i=1}^{2n} X_i^2 + \sum_{i=1}^{n} X_{2i-1}X_{2i} = Z^T \Delta Z$

由 Δ 为正定矩阵（$2n$ 阶）知，$Y = Z^T \Delta Z \sim \chi^2_{2n}$

11. 令 $X_1, X_2, \cdots, X_{2n}$ $iid \sim N(0,1)$. 求

$$Y = \frac{1}{2}\sum_{i=1}^{2n} X_i^2 + \sum_{i=1}^{n} X_{2i-1}X_{2i}$$

的分布。

令 $X = (X_1, X_2, \cdots, X_{2n})^T$ 则 $X \sim N(0_n, I_n)$.

$Y = \frac{1}{2} X^T A X$ 其中 $A = \begin{pmatrix} 1 & 1 & & & & \\ 1 & 1 & & & & \\ & & 1 & 1 & & \\ & & 1 & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \quad 1 \\ & & & & & 1 \quad 1 \end{pmatrix}_{2n \times 2n}$

$\mathrm{rank}(A) = n$ 由 χ^2 分布的性质与，$Y \sim \chi^2_n$

2. 记 $X_1, X_2, \cdots, X_n$ iid $\sim N(0,1)$,

$$Y = a\left(X_1 + X_2 + \cdots + X_m\right)^2 + b\left(X_{m+1} + X_{m+2} + \cdots + X_n\right)^2$$
$$Z = \frac{c\left(X_1 + X_2 + \cdots + X_m\right)}{\sqrt{X_{m+1}^2 + X_{m+2}^2 + \cdots + X_n^2}}, m < n$$

若 Y 服从 χ^2 分布, Z 服从 t 分布, 求 a, b, c.

Ans: (1) 记

$$Z_1 = X_1 + X_2 + \cdots + X_m \sim N(0, m),$$
$$Z_2 = X_{m+1} + X_{m+2} + \cdots + X_n \sim N(0, n-m),$$

且 Z_1 与 Z_2 相互独立. 由于

$$\frac{Z_1}{\sqrt{m}} \sim N(0,1), \quad \frac{Z_2}{\sqrt{n-m}} \sim N(0,1),$$

故

$$\left(\frac{Z_1}{\sqrt{m}}\right)^2 + \left(\frac{Z_2}{\sqrt{n-m}}\right)^2 = \frac{1}{m} z_1^2 + \frac{1}{n-m} z_2^2 \sim \chi_2^2.$$

此时, 常数 $a = \dfrac{1}{m}, b = \dfrac{1}{n-m}$.

(2) 记

$$Z_3 = \chi_{m+1}^2 + \chi_{m+2}^2 + \cdots + \chi_n^2 \sim \chi_{n-m}^2.$$

由 t 分布定义,

$$\frac{\frac{Z_1}{\sqrt{m}}}{\sqrt{Z_3/(n-m)}} = \sqrt{\frac{n-m}{m}}, \quad \frac{Z_1}{\sqrt{Z_3}} \sim t_{n-m}.$$

从而 $c = \sqrt{\dfrac{n-m}{m}}$.

3. 记 $X_1, X_2, \cdots, X_n$ iid $\sim N(0,1)$, 样本方差为 S^2. 求 c 使得 $\dfrac{c\bar{X}^2}{S^2}$ 服从 F 分布.

Ans: 由 $\bar{X} \sim N\left(0, \frac{1}{n}\right)$, 知 $\left(\bar{X} / \sqrt{\frac{1}{n}}\right)^2 \sim \chi_1^2$. 根据 F 分布的定义, 有

$$\frac{\left(\bar{X} / \sqrt{\frac{1}{n}}\right)^2}{(n-1)S^2/(n-1)} = \frac{n\bar{X}^2}{S^2} \sim F_{1,n-1}.$$

故当 $c = n$ 时, 有 $\frac{C\bar{X}^2}{S^2} \sim F_{1,n-1}$.

4. 求: (1) $X_{n+1} + X_{n+2} - \frac{1}{n+1} \sum_{i=1}^{n+1} X_i$ 的分布;

(2) $\frac{X_{n+1}^2 + X_{n+2}^2}{2S^2}$ 的分布;

(3) a, b 使得 $\frac{a\left(X_{n+1} + X_{n+2} - \bar{X}\right)}{S}, \frac{b\left(X_{n+1} - X_{n+2} - \bar{X}\right)}{S}$ 都服从 t 分布, 并给出他们的自由度.

Ans: (1)

$$X_{n+1} + X_{n+2} - \frac{1}{n+1} \sum_{i=1}^{n+1} X_i = \frac{n}{n+1} X_{m+1} + X_{m+2} - \frac{1}{n+1} \sum_{i=1}^{n} X_i$$

由 $X_1, X_2, \ldots, X_n, X_{m+1}, X_{m+2}$ iid $\sim N\left(0, \sigma^2\right)$, 则

$$\frac{n}{n+1} X_{n+1} + X_{n+2} - \frac{1}{n+1} \sum_{n}^{i=1} X_i \sim N\left(0, \sigma_n^2\right),$$

其中

$$\begin{aligned}
\sigma_*^2 &= \operatorname{Var}\left(\frac{n}{n^2} X_{n+1} + X_{n+2} - \frac{1}{n+1} \sum_{i=1}^{n} X_i\right) \\
&= \left(\frac{-n}{n+1}\right)^2 \cdot \operatorname{Var}\left(X_{n+1}\right) + \operatorname{Var}\left(X_{n+2}\right) + \left(\frac{1}{n+1}\right)^2 \cdot \sum_{i=1}^{n} \operatorname{Var}\left(X_i\right) \\
&= \left(\frac{n^2+n}{(n+1)^2} + 1\right) \sigma^2 \\
&= \frac{2n+1}{n+1} \sigma^2.
\end{aligned}$$

(2) 由题意

$$\left(\frac{X_{n+1}}{\sigma}\right)^2+\left(\frac{X_{n+2}}{\sigma}\right)^2\sim\chi_2^2$$

$$\frac{(n-1)S^2}{\sigma^2}\sim\chi_{n-1}^2.$$

且 X_{n+1},X_{n+2},S^2 相互独立，由 F 分布的构造

$$\frac{\left[\left(\frac{x_{n+1}}{\sigma}\right)^2+\left(\frac{x_{n+2}}{\sigma}\right)^2\right]/2}{\frac{(n-1)s^2}{\sigma^2}/(n-1)}=\frac{X_{n+1}^2+X_{n+2}^2}{2S^2}\sim F_{2n-1+}$$

(3) 由

$$X_{m+1}+X_{n+2}-\hat{X}\sim N\left(0,\left(2+\frac{1}{n}\right)\sigma^2\right)$$

$$\frac{(n-1)S^2}{\sigma^2}\sim X_{n-1}^2$$

且 $X_{n+1}+X_{n+2}-\hat{X}$ 与 S^2 相互独立, 有

$$\frac{\left(X_{n+1}+X_{n+2}-\bar{X}\right)/\sqrt{\left(2+\frac{1}{n}\right)\sigma^2}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}}=\frac{\sqrt{\frac{n}{n+1}}\left(X_{m+1}+X_{n+2}-\bar{X}\right)}{S}\sim t_{n-1}$$

故 $a=\sqrt{\frac{n}{2n+1}}$, 自由度为 $n-1$.

同理,

$$X_{n+1}-X_{n+2}-\bar{X}\sim N\left(0,\left(2+\frac{1}{n}\right)\sigma^2\right)$$

有 $b=\sqrt{\frac{n}{2n+1}}$, 使

$$\frac{b\left(X_{n+1}-X_{n+2}-\bar{X}\right)}{S}\sim t_{n-1}$$

自由度为 $n-1$.

5. 记 $(X_1,X_2,\cdots,X_k)^{\mathrm{T}}\sim\mathrm{Mul}\,(N,p_1,p_2,\cdots,p_k)$. 求 $p_1,p_2,\cdots,p_k$ 的矩估计和最大似然估计.

Ans: (1) X_i 的边缘分布为二项分布 $B(N, p_i)$, $\quad i=1,2,\ldots,k$, 则 $P(X_i = N_i) = \binom{N}{N_i} p_i^{N_i} (1-p_i)^{N-N_i}$, $EX_i = Np_i$.

令 $N_i = Np_i \Rightarrow p_i = \dfrac{N_i}{N}, \quad i=1,2,\ldots,k.$

从而 p_i 的矩估计为 $\hat{p}_i = \dfrac{N_i}{N}, \quad i=1,2,\ldots,k.$

(2) 似然函数为:

$$L(p_1, p_2, \ldots, p_k) = \frac{N!}{N_1! N_2! \cdots N_k!} p_1^{N_1} p_2^{N_2} \cdots p_k^{N_k}$$

对数似然函数为:

$$l(p_1, p_2, \ldots, p_k) = \ln(N!) - \sum_{i=1}^{k} \ln(N_i) + \sum_{i=1}^{k} N_i \ln(p_i)$$

其中 $\sum_{i=1}^{k} p_i = 1$.

用 lagrange 乘子法求有约束条件的极值问题:

记 $\psi(p_1, p_2, \ldots, p_k, \lambda) = l(p_1, P_2, \ldots, p_k) - \lambda\left(\sum_{i=1}^{k} p_i - 1\right)$,

关于 p_i 求导, 有 $\dfrac{\partial \psi(p_1, P_2, \ldots, p_k, \lambda)}{\partial p_i} = \dfrac{N_i}{p_i} - \lambda$,

令上式为 0 , 得 $p_i = \dfrac{N_i}{\lambda}$. 由 $\sum_{i=1}^{k} p_i = \sum_{i=1}^{k} \dfrac{N_i}{\lambda} = \dfrac{N}{\lambda} = 1 \Rightarrow \lambda = N$.

进而 $\hat{p}_i = \dfrac{N_i}{N}$ 是 p_i 的极大似然估计 $i=1,2,\ldots,k$.

6. n 个人中拥有上述三种基因型的人数分别为 n_0, n_1, n_2. 求 p 的矩估计和最大似然估计.

Ans: 记 n 个人基因型含有 G 的个数为 $X_1, X_2, \ldots, X_n$, 他们独立同分布, 且概率函数为

取值	0	1	2
概率	$(1-p)^2$	$2p(1-p)$	p^2

(1) 记 $n_j = \sum_{j=1}^{n} I(X_j = j), j = 0,1,2,$

$(n_0, n_1, n_2)^T \sim \mathrm{Mul}(n; (1-p)^2, 2p(1-p), p^2)$, $n_0 \sim B(n, (1-p)^2), n_1 \sim B(n, 2p(1-$

$p)), n_2 \sim B\left(n, p^2\right)$, 设 $n_0 = n \cdot (1-p)^2$ 或 $n_1 = n \cdot 2p(1-p)$ 或 $n_2 = n \cdot p^2$, 得 $p = 1 - \sqrt{\frac{n_0}{n}}$ 或 $p = \pm\sqrt{\frac{1}{4} - \frac{n_1}{2n}} + \frac{1}{2}$ 或 $p = \sqrt{\frac{n_2}{n}}$. 均为 p 的矩估计.

(2) 似然函数为

$$L(p) = \left[(1-p)^2\right]^{n_0} \left[2p(1-p)\right]^{n_1} \left[p^2\right]^{n_2}$$

对数似然函数为

$$l(p) = n_1 \ln 2 + (n_1 + 2n_2) \ln p + (2n_0 + n_1) \ln(1-p)$$

似然方程为

$$\frac{\partial l(p)}{\partial p} = \frac{(n_1 + 2n_2)}{p} - \frac{(2n_0 + n_1)}{1-p} \Rightarrow p = \frac{n_1 + 2n_2}{2\left(n_0 + n_1 + n_2\right)} = \frac{n_1 + 2n_2}{2n}$$

p 的极大似然估计为 $\hat{p} = \frac{n_1 + 2n_2}{2n}$.

这次作业没有补充材料，点估计和极大似然估计作为数理统计中最重要的方法之一，希望大家自行学习网上的众多资料，而不局限于我们给出的补充材料.