

数理统计

第二章

抽样分布及若干预备知识

2026 年 3 月 25 日

- ① 2.4 次序统计量的分布
 - 2.4.1 次序统计量的分布
 - 2.4.2 极差的分布
 - 2.4.3 均匀分布的情形

2.4 次序统计量的分布

若 X_1, X_2, \dots, X_n *i.i.d.* $\sim F$ ，将其按大小排列为 $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ ，则称 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 (X_1, X_2, \dots, X_n) 的次序统计量。

2.4 次序统计量的分布

若 X_1, X_2, \dots, X_n *i.i.d.* $\sim F$ ，将其按大小排列

为 $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ ，则称 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 (X_1, X_2, \dots, X_n) 的**次序统计量**。

它的任何一部分，如 $X_{(i)}$ 和 $(X_{(i)}, X_{(j)})$ ($1 \leq i < j \leq n$)等也称为次序统计量。

2.4 次序统计量的分布

若 X_1, X_2, \dots, X_n *i.i.d.* $\sim F$ ，将其按大小排列为 $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ ，则称 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 (X_1, X_2, \dots, X_n) 的**次序统计量**。

它的任何一部分，如 $X_{(i)}$ 和 $(X_{(i)}, X_{(j)})$ ($1 \leq i < j \leq n$)等也称为次序统计量。

本节将导出在 F 有密度时**单个次序统计量的分布**、**次序统计量的联合分布**和**极差的分布**。

2.4.1 次序统计量的分布

引理 (2.4.1)

设总体 X 的分布函数为 $F(x)$ ，密度函数为 $f(x)$ ， X_1, \dots, X_n 为从总体 X 中抽取的简单样本，则有

$$\int_{a < x_1 < \dots, x_n < b} f(x_1) \cdots f(x_n) dx_1 \cdots dx_n = \frac{1}{n!} [F(b) - F(a)]^n,$$

此处 a 可取有限数或 $-\infty$ ， b 可取有限数或 ∞ 。

2.4.1 次序统计量的分布

证

采用数学归纳法证明。

2.4.1 次序统计量的分布

证

采用数学归纳法证明。

当 $n = 2$ 时有,

2.4.1 次序统计量的分布

证

采用数学归纳法证明。

当 $n = 2$ 时有,

$$\iint_{a < x_1 < x_2 < b} f(x_1)f(x_2)dx_1dx_2$$

2.4.1 次序统计量的分布

证

采用数学归纳法证明。

当 $n = 2$ 时有,

$$\iint_{a < x_1 < x_2 < b} f(x_1)f(x_2)dx_1dx_2 = \int_a^b f(x_1) \int_{x_1}^b f(x_2)dx_2dx_1$$

=

2.4.1 次序统计量的分布

证

采用数学归纳法证明。

当 $n = 2$ 时有,

$$\begin{aligned} & \iint_{a < x_1 < x_2 < b} f(x_1)f(x_2)dx_1dx_2 = \int_a^b f(x_1) \int_{x_1}^b f(x_2)dx_2dx_1 \\ &= \int_a^b f(x_1)[F(b) - F(x_1)]dx_1 = F(b)[F(b) - F(a)] - \int_a^b F(x_1)dF(x_1) \\ &= \end{aligned}$$

2.4.1 次序统计量的分布

证

采用数学归纳法证明。

当 $n = 2$ 时有,

$$\begin{aligned} \iint_{a < x_1 < x_2 < b} f(x_1)f(x_2)dx_1dx_2 &= \int_a^b f(x_1) \int_{x_1}^b f(x_2)dx_2dx_1 \\ &= \int_a^b f(x_1)[F(b) - F(x_1)]dx_1 = F(b)[F(b) - F(a)] - \int_a^b F(x_1)dF(x_1) \\ &= F(b)[F(b) - F(a)] - \frac{[F(x_1)]^2}{2} \Big|_a^b = \frac{[F(b) - F(a)]^2}{2}, \end{aligned}$$

等式成立。

2.4.1 次序统计量的分布

设当 $n = k$ 时等式成立，即

2.4.1 次序统计量的分布

设当 $n = k$ 时等式成立, 即

$$\int_{a < x_1 < \dots < x_k < b} f(x_1) \cdots f(x_k) dx_1 \cdots dx_k = \frac{1}{k!} [F(b) - F(a)]^k.$$

2.4.1 次序统计量的分布

设当 $n = k$ 时等式成立, 即

$$\int_{a < x_1 < \dots < x_k < b} f(x_1) \cdots f(x_k) dx_1 \cdots dx_k = \frac{1}{k!} [F(b) - F(a)]^k.$$

则当 $n = k + 1$ 时有

$$\int_{a < x_1 < \dots < x_k < x_{k+1} < b} f(x_1) \cdots f(x_k) f(x_{k+1}) dx_1 \cdots dx_k dx_{k+1}$$

=

2.4.1 次序统计量的分布

设当 $n = k$ 时等式成立, 即

$$\int_{a < x_1 < \dots < x_k < b} f(x_1) \cdots f(x_k) dx_1 \cdots dx_k = \frac{1}{k!} [F(b) - F(a)]^k.$$

则当 $n = k + 1$ 时有

$$\begin{aligned} & \int_{a < x_1 < \dots < x_k < x_{k+1} < b} f(x_1) \cdots f(x_k) f(x_{k+1}) dx_1 \cdots dx_k dx_{k+1} \\ &= \int_a^b f(x_{k+1}) \int_{a < x_1 < \dots < x_k < x_{k+1}} f(x_1) \cdots f(x_k) dx_1 \cdots dx_k dx_{k+1} \\ &= \end{aligned}$$

2.4.1 次序统计量的分布

设当 $n = k$ 时等式成立, 即

$$\int_{a < x_1 < \dots < x_k < b} f(x_1) \cdots f(x_k) dx_1 \cdots dx_k = \frac{1}{k!} [F(b) - F(a)]^k.$$

则当 $n = k + 1$ 时有

$$\begin{aligned} & \int_{a < x_1 < \dots < x_k < x_{k+1} < b} f(x_1) \cdots f(x_k) f(x_{k+1}) dx_1 \cdots dx_k dx_{k+1} \\ &= \int_a^b f(x_{k+1}) \int_{a < x_1 < \dots < x_k < x_{k+1}} f(x_1) \cdots f(x_k) dx_1 \cdots dx_k dx_{k+1} \\ &= \int_a^b f(x_{k+1}) \frac{1}{k!} [F(x_{k+1}) - F(a)]^k dx_{k+1} \end{aligned}$$

2.4.1 次序统计量的分布

$$= \int_a^b \frac{1}{k!} [F(x_{k+1}) - F(a)]^k d[F(x_{k+1}) - F(a)]$$
$$=$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \int_a^b \frac{1}{k!} [F(x_{k+1}) - F(a)]^k d[F(x_{k+1}) - F(a)] \\ &= \frac{[F(x_{k+1}) - F(a)]^{k+1}}{(k+1)!} \Big|_a^b = \frac{[F(b) - F(a)]^{k+1}}{(k+1)!}, \end{aligned}$$

等式成立。这就证明了引理。

□

2.4.1 次序统计量的分布

1. 次序统计量 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 的联合分布

2.4.1 次序统计量的分布

1. 次序统计量 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 的联合分布

定理 (2.4.1)

设总体 X 有密度函数 $f(x)$, $-\infty < x < \infty$, X_1, X_2, \dots, X_n 为从总体 X 中抽取的简单样本, 如前所述 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 (X_1, X_2, \dots, X_n) 的次序统计量。令 $Y_i = X_{(i)}$, $i = 1, \dots, n$, 则次序统计量 (Y_1, Y_2, \dots, Y_n) 的联合密度为

2.4.1 次序统计量的分布

1. 次序统计量 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 的联合分布

定理 (2.4.1)

设总体 X 有密度函数 $f(x)$, $-\infty < x < \infty$, X_1, X_2, \dots, X_n 为从总体 X 中抽取的简单样本, 如前所述 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 (X_1, X_2, \dots, X_n) 的次序统计量。令 $Y_i = X_{(i)}$, $i = 1, \dots, n$, 则次序统计量 (Y_1, Y_2, \dots, Y_n) 的联合密度为

$$g(y_1, y_2, \dots, y_n) = \begin{cases} n! f(y_1) f(y_2) \dots f(y_n), & y_1 < y_2 < \dots < y_n, \\ 0, & \text{其他.} \end{cases}$$

2.4.1 次序统计量的分布

证

作下列变换：

2.4.1 次序统计量的分布

证

作下列变换：

$$\left\{ \begin{array}{l} y_1 = x_{i_1}, \\ y_2 = x_{i_2}, \\ \dots\dots\dots \text{这里 } x_{i_1} < x_{i_2} < \dots < x_{i_n}, \\ y_n = x_{i_n}, \end{array} \right.$$

2.4.1 次序统计量的分布

证

作下列变换：

$$\left\{ \begin{array}{l} y_1 = x_{i_1}, \\ y_2 = x_{i_2}, \\ \dots\dots\dots \text{这里 } x_{i_1} < x_{i_2} < \dots < x_{i_n}, \\ y_n = x_{i_n}, \end{array} \right.$$

设 (i_1, i_2, \dots, i_n) 是 $(1, 2, \dots, n)$ 的任一排列，这样的排列共有 $n!$ 种。

2.4.1 次序统计量的分布

证

作下列变换：

$$\begin{cases} y_1 = x_{i_1}, \\ y_2 = x_{i_2}, \\ \dots\dots\dots \text{这里 } x_{i_1} < x_{i_2} < \dots < x_{i_n}, \\ y_n = x_{i_n}, \end{cases}$$

设 (i_1, i_2, \dots, i_n) 是 $(1, 2, \dots, n)$ 的任一排列，这样的排列共有 $n!$ 种。

因此上述一一对应的变换有 $n!$ 个。

2.4.1 次序统计量的分布

与上述 $n!$ 个变换相对应的 $n!$ 个区域 $x_{i_1} < x_{i_2} < \cdots < x_{i_n}$ 是互不相交的。

2.4.1 次序统计量的分布

与上述 $n!$ 个变换相对应的 $n!$ 个区域 $x_{i_1} < x_{i_2} < \cdots < x_{i_n}$ 是互不相交的。

这 $n!$ 组区域和某些概率为零的集合（如 $x_1 = x_2 = \cdots = x_n$ 等）的并构成 n 维欧式空间 \mathcal{R}^n 。

2.4.1 次序统计量的分布

与上述 $n!$ 个变换相对应的 $n!$ 个区域 $x_{i_1} < x_{i_2} < \cdots < x_{i_n}$ 是互不相交的。

这 $n!$ 组区域和某些概率为零的集合（如 $x_1 = x_2 = \cdots = x_n$ 等）的并构成 n 维欧式空间 \mathcal{R}^n 。

每一组变换的Jacobi行列式

$$|J| = \left| \frac{\partial(x_{i_1}, x_{i_2}, \dots, x_{i_n})}{\partial(y_1, y_2, \dots, y_n)} \right| =$$

2.4.1 次序统计量的分布

与上述 $n!$ 个变换相对应的 $n!$ 个区域 $x_{i_1} < x_{i_2} < \cdots < x_{i_n}$ 是互不相交的。

这 $n!$ 组区域和某些概率为零的集合（如 $x_1 = x_2 = \cdots = x_n$ 等）的并构成 n 维欧氏空间 \mathcal{R}^n 。

每一组变换的Jacobi行列式

$$|J| = \left| \frac{\partial(x_{i_1}, x_{i_2}, \dots, x_{i_n})}{\partial(y_1, y_2, \dots, y_n)} \right| = \begin{vmatrix} \frac{\partial x_{i_1}}{\partial y_1} & \frac{\partial x_{i_1}}{\partial y_2} & \cdots & \frac{\partial x_{i_1}}{\partial y_n} \\ \frac{\partial x_{i_2}}{\partial y_1} & \frac{\partial x_{i_2}}{\partial y_2} & \cdots & \frac{\partial x_{i_2}}{\partial y_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial x_{i_n}}{\partial y_1} & \frac{\partial x_{i_n}}{\partial y_2} & \cdots & \frac{\partial x_{i_n}}{\partial y_n} \end{vmatrix}$$

2.4.1 次序统计量的分布

$$= \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

2.4.1 次序统计量的分布

$$= \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

由于 $X_{i_1}, X_{i_2}, \dots, X_{i_n}$ 是总体 X 的*i.i.d.*样本, 故 $(X_{i_1}, X_{i_2}, \dots, X_{i_n})$ 的联合密度为

2.4.1 次序统计量的分布

$$= \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

由于 $X_{i_1}, X_{i_2}, \dots, X_{i_n}$ 是总体 X 的*i.i.d.*样本, 故 $(X_{i_1}, X_{i_2}, \dots, X_{i_n})$ 的联合密度为

$$\begin{aligned} h(x_{i_1}, x_{i_2}, \dots, x_{i_n}) &= \prod_{j=1}^n f(x_{i_j}) I_{\{x_{i_1} < x_{i_2} < \cdots < x_{i_n}\}} \\ &= f(x_1) f(x_2) \cdots f(x_n) = \prod_{i=1}^n f(x_i), \end{aligned}$$

2.4.1 次序统计量的分布

在上述变换下, (Y_1, Y_2, \dots, Y_n) 的联合密度为

2.4.1 次序统计量的分布

在上述变换下, (Y_1, Y_2, \dots, Y_n) 的联合密度为

$$g(y_1, y_2, \dots, y_n) \\ = \sum_{\begin{pmatrix} 1, 2, \dots, n \\ i_1, i_2, \dots, i_n \end{pmatrix}} |J| \cdot h(x_{i_1}(y_1, \dots, y_n), x_{i_2}(y_1, \dots, y_n), \dots, x_{i_n}(y_1, \dots, y_n))$$

2.4.1 次序统计量的分布

在上述变换下, (Y_1, Y_2, \dots, Y_n) 的联合密度为

$$\begin{aligned} & g(y_1, y_2, \dots, y_n) \\ = & \sum_{\binom{1, 2, \dots, n}{i_1, i_2, \dots, i_n}} |J| \cdot h(x_{i_1}(y_1, \dots, y_n), x_{i_2}(y_1, \dots, y_n), \dots, x_{i_n}(y_1, \dots, y_n)) \\ = & \begin{cases} n! f(y_1) f(y_2) \dots f(y_n), & y_1 < y_2 < \dots < y_n, \\ 0, & \text{其他,} \end{cases} \end{aligned}$$

2.4.1 次序统计量的分布

此处

$$\sum_{\binom{1, 2, \dots, n}{i_1, i_2, \dots, i_n}}$$

表示对 i_1, i_2, \dots, i_n 遍历 $1, 2, \dots, n$ 的所有排列求和（共 $n!$ 种）。 □

2.4.1 次序统计量的分布

2. 两个次序统计量($X_{(i)}, X_{(j)}$)的联合分布

2.4.1 次序统计量的分布

2. 两个次序统计量 $(X_{(i)}, X_{(j)})$ 的联合分布

定理 (2.4.2)

设 $X_1, X_2, \dots, X_n, i.i.d. \sim F$, F 有密度 f 。记 $\mathbf{X} = (X_1, \dots, X_n)$, 则 \mathbf{X} 的次序统计量 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 中的任意两个 $(X_{(i)}, X_{(j)})$ ($i < j$) 的联合密度为

2.4.1 次序统计量的分布

2. 两个次序统计量 $(X_{(i)}, X_{(j)})$ 的联合分布

定理 (2.4.2)

设 $X_1, X_2, \dots, X_n, i.i.d. \sim F$, F 有密度 f 。记 $\mathbf{X} = (X_1, \dots, X_n)$, 则 \mathbf{X} 的次序统计量 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 中的任意两个 $(X_{(i)}, X_{(j)})$ ($i < j$) 的联合密度为

$$f_{ij}(x, y) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} \\ \quad \times [1 - F(y)]^{n-j} f(x) f(y), & x < y, \\ 0, & \text{其他.} \end{cases} \quad (1)$$

2.4.1 次序统计量的分布

证

已知次序统计量 $(X_{(1)}, \dots, X_{(n)})$ 的联合密度函数为

2.4.1 次序统计量的分布

证

已知次序统计量 $(X_{(1)}, \dots, X_{(n)})$ 的联合密度函数为

$$g_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = n! f(x_1) \cdots f(x_n), \quad x_1 < x_2 < \cdots < x_n.$$

2.4.1 次序统计量的分布

证

已知次序统计量 $(X_{(1)}, \dots, X_{(n)})$ 的联合密度函数为

$$g_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = n! f(x_1) \cdots f(x_n), \quad x_1 < x_2 < \cdots < x_n.$$

当 $x < y$ 时, 任意两个次序统计量 $(X_{(i)}, X_{(j)})$, $(i < j)$ 的联合密度为

2.4.1 次序统计量的分布

证

已知次序统计量 $(X_{(1)}, \dots, X_{(n)})$ 的联合密度函数为

$$g_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) = n! f(x_1) \cdots f(x_n), \quad x_1 < x_2 < \cdots < x_n.$$

当 $x < y$ 时, 任意两个次序统计量 $(X_{(i)}, X_{(j)})$, $(i < j)$ 的联合密度为

$$f_{ij}(x, y) = \int_{-\infty < x_1 < \cdots < x_n < \infty} g_{X_{(1)}, \dots, X_{(n)}}(x_1, \dots, x_n) dx_1 \cdots dx_{i-1} \\ dx_{i+1} \cdots dx_{j-1} dx_{j+1} \cdots dx_n,$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \int_{-\infty < x_1 < \dots < x_n < \infty} \dots \int n! f(x_1) \dots f(x_n) dx_1 \dots dx_{i-1} \\ &\quad dx_{i+1} \dots dx_{j-1} dx_{j+1} \dots dx_n, \\ &= \end{aligned}$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \int_{-\infty < x_1 < \dots < x_n < \infty} \dots \int n! f(x_1) \dots f(x_n) dx_1 \dots dx_{i-1} \\ &\quad dx_{i+1} \dots dx_{j-1} dx_{j+1} \dots dx_n, \\ &= n! \int_{-\infty < x_1 < \dots < x_{i-1} < x} \dots \int f(x_1) \dots f(x_{i-1}) dx_1 \dots dx_{i-1} \times f(x) \end{aligned}$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \int_{-\infty < x_1 < \dots < x_n < \infty} \dots \int n! f(x_1) \dots f(x_n) dx_1 \dots dx_{i-1} \\ &\quad dx_{i+1} \dots dx_{j-1} dx_{j+1} \dots dx_n, \\ &= n! \int_{-\infty < x_1 < \dots < x_{i-1} < x} \dots \int f(x_1) \dots f(x_{i-1}) dx_1 \dots dx_{i-1} \times f(x) \\ &\quad \times \int_{x < x_{i+1} < \dots < x_{j-1} < y} \dots \int f(x_{i+1}) \dots f(x_{j-1}) dx_{i+1} \dots dx_{j-1} \times f(y) \end{aligned}$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \int_{-\infty < x_1 < \dots < x_n < \infty} \cdots \int n! f(x_1) \dots f(x_n) dx_1 \cdots dx_{i-1} \\ &\quad dx_{i+1} \cdots dx_{j-1} dx_{j+1} \cdots dx_n, \\ &= n! \int_{-\infty < x_1 < \dots < x_{i-1} < x} \cdots \int f(x_1) \dots f(x_{i-1}) dx_1 \cdots dx_{i-1} \times f(x) \\ &\quad \times \int_{x < x_{i+1} < \dots < x_{j-1} < y} \cdots \int f(x_{i+1}) \dots f(x_{j-1}) dx_{i+1} \cdots dx_{j-1} \times f(y) \\ &\quad \times \int_{y < x_{j+1} < \dots < x_n < \infty} \cdots \int f(x_{j+1}) \dots f(x_n) dx_{j+1} \cdots dx_n \end{aligned}$$

2.4.1 次序统计量的分布

因为 $F(\infty) = 1$, $F(-\infty) = 0$, 由引理2.4.1可知, 上式等于

=

2.4.1 次序统计量的分布

因为 $F(\infty) = 1$, $F(-\infty) = 0$, 由引理2.4.1可知, 上式等于

$$\begin{aligned} &= n! \frac{f(x)}{(i-1)!} [F(x) - F(-\infty)]^{i-1} \times \frac{f(y)}{(j-i-1)!} [F(y) - F(x)]^{j-i-1} \\ &\quad \times \frac{[F(\infty) - F(y)]^{n-j}}{(n-j)!} \\ &= \end{aligned}$$

2.4.1 次序统计量的分布

因为 $F(\infty) = 1$, $F(-\infty) = 0$, 由引理2.4.1可知, 上式等于

$$\begin{aligned} &= n! \frac{f(x)}{(i-1)!} [F(x) - F(-\infty)]^{i-1} \times \frac{f(y)}{(j-i-1)!} [F(y) - F(x)]^{j-i-1} \\ &\quad \times \frac{[F(\infty) - F(y)]^{n-j}}{(n-j)!} \\ &= \frac{n! f(x) f(y)}{(i-1)! (j-i-1)! (n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} [1 - F(y)]^{n-j}. \end{aligned}$$

2.4.1 次序统计量的分布

因为 $F(\infty) = 1$, $F(-\infty) = 0$, 由引理2.4.1可知, 上式等于

$$\begin{aligned} &= n! \frac{f(x)}{(i-1)!} [F(x) - F(-\infty)]^{i-1} \times \frac{f(y)}{(j-i-1)!} [F(y) - F(x)]^{j-i-1} \\ &\quad \times \frac{[F(\infty) - F(y)]^{n-j}}{(n-j)!} \\ &= \frac{n! f(x) f(y)}{(i-1)! (j-i-1)! (n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} [1 - F(y)]^{n-j}. \end{aligned}$$

当 $x \geq y$ 时, $f_{ij}(x, y) = 0$, 这就证明了定理。 □

2.4.1 次序统计量的分布

用类似的方法可求 r ($r > 2$)个次序统计量 $(X_{(i_1)}, \dots, X_{(i_r)})$
($i_1 < \dots < i_r$) 的联合分布。

2.4.1 次序统计量的分布

注 2.4.1: 启发式推导, 求次序统计量分布的另一种方法

2.4.1 次序统计量的分布

注 2.4.1: 启发式推导, 求次序统计量分布的另一种方法

在 n 次独立的试验中, 事件 $x < X_{(i)} < x + \Delta_i, y < X_{(j)} < y + \Delta_j$ 的概率为

2.4.1 次序统计量的分布

注 2.4.1: 启发式推导, 求次序统计量分布的另一种方法

在 n 次独立的试验中, 事件 $x < X_{(i)} < x + \Delta_i$, $y < X_{(j)} < y + \Delta_j$ 的概率为

$$\begin{aligned} & P(x < X_{(i)} < x + \Delta_i, y < X_{(j)} < y + \Delta_j) \\ &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(x + \Delta_i) - F(x)] \\ & \quad \times [F(y) - F(x + \Delta_i)]^{j-i-1} [F(y + \Delta_j) - F(y)] [1 - F(y + \Delta_j)]^{n-j} \end{aligned}$$

2.4.1 次序统计量的分布

对上式两端同除 $\Delta_i \Delta_j$ 并令其趋近于零，可得到

2.4.1 次序统计量的分布

对上式两端同除 $\Delta_i\Delta_j$ 并令其趋近于零, 可得到

$$f_{ij}(x, y) = \lim_{\Delta_i \rightarrow 0, \Delta_j \rightarrow 0} \frac{P(x < X_{(i)} < x + \Delta_i, y < X_{(j)} < y + \Delta_j)}{\Delta_i \Delta_j}$$

=

2.4.1 次序统计量的分布

对上式两端同除 $\Delta_i\Delta_j$ 并令其趋近于零, 可得到

$$\begin{aligned} f_{ij}(x, y) &= \lim_{\Delta_i \rightarrow 0, \Delta_j \rightarrow 0} \frac{P(x < X_{(i)} < x + \Delta_i, y < X_{(j)} < y + \Delta_j)}{\Delta_i \Delta_j} \\ &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} \\ &\quad \times \lim_{\Delta_i \rightarrow 0} \frac{[F(x + \Delta_i) - F(x)]}{\Delta_i} \\ &\quad \times \lim_{\Delta_i \rightarrow 0} [F(y) - F(x + \Delta_i)]^{j-i-1} \\ &\quad \times \lim_{\Delta_j \rightarrow 0} \frac{[F(y + \Delta_j) - F(y)]}{\Delta_j} \\ &\quad \times \lim_{\Delta_j \rightarrow 0} [1 - F(y + \Delta_j)]^{n-j} \end{aligned}$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} \\ &\quad \times f(x)[F(y) - F(x)]^{j-i-1} f(y)[1 - F(y)]^{n-j}. \end{aligned}$$

□

2.4.1 次序统计量的分布

3. 单个次序统计量 $X_{(m)}$ 的分布

定理 (2.4.3)

设总体 X 有密度函数 $f(x)$, $-\infty < x < \infty$, 令 X_1, X_2, \dots, X_n 为从总体 X 中抽取的简单样本, 如本节开头所述 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 (X_1, X_2, \dots, X_n) 的次序统计量, 则第 m 个次序统计量 $X_{(m)}$ 概率密度函数为

2.4.1 次序统计量的分布

3. 单个次序统计量 $X_{(m)}$ 的分布

定理 (2.4.3)

设总体 X 有密度函数 $f(x)$, $-\infty < x < \infty$, 令 X_1, X_2, \dots, X_n 为从总体 X 中抽取的简单样本, 如本节开头所述 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 (X_1, X_2, \dots, X_n) 的次序统计量, 则第 m 个次序统计量 $X_{(m)}$ 概率密度函数为

$$f_m(x) = m \binom{n}{m} [F(x)]^{m-1} [1 - F(x)]^{n-m} f(x). \quad (2)$$

此处 $x = x_{(m)}$ 。

2.4.1 次序统计量的分布

定理 (2.4.3续)

其分布函数为

$$F_m(x) = m \binom{n}{m} \int_0^{F(x)} t^{m-1} (1-t)^{n-m} dt.$$

2.4.1 次序统计量的分布

定理 (2.4.3续)

其分布函数为

$$F_m(x) = m \binom{n}{m} \int_0^{F(x)} t^{m-1} (1-t)^{n-m} dt.$$

特别当取 $m = 1$ 时得到样本极小值 $X_{(1)}$ 的密度函数为

$$f_1(x) = n[1 - F(x)]^{n-1} f(x).$$

2.4.1 次序统计量的分布

定理 (2.4.3续)

当取 $m = n$ 时得到样本极大值 $X_{(n)}$ 的密度函数为

2.4.1 次序统计量的分布

定理 (2.4.3续)

当取 $m = n$ 时得到样本极大值 $X_{(n)}$ 的密度函数为

$$f_n(x) = n[F(x)]^{n-1}f(x).$$

2.4.1 次序统计量的分布

证

由于 $f_m(x)$ 是 $(X_{(1)}, \dots, X_{(n)})$ 的联合密度 $h(x_{(1)}, \dots, x_{(n)})$ 的边缘密度, 由引理2.4.1和定理2.4.1可知

2.4.1 次序统计量的分布

证

由于 $f_m(x)$ 是 $(X_{(1)}, \dots, X_{(n)})$ 的联合密度 $h(x_{(1)}, \dots, x_{(n)})$ 的边缘密度, 由引理2.4.1和定理2.4.1可知

$$\begin{aligned} f_m(x) &= \int \cdots \int_{-\infty < x_1 < \cdots < x_n < \infty} n! f(x_1) \cdots f(x_n) dx_1 \cdots dx_{m-1} \\ &\quad dx_{m+1} \cdots dx_n \\ &= \end{aligned}$$

2.4.1 次序统计量的分布

证

由于 $f_m(x)$ 是 $(X_{(1)}, \dots, X_{(n)})$ 的联合密度 $h(x_{(1)}, \dots, x_{(n)})$ 的边缘密度, 由引理2.4.1和定理2.4.1可知

$$\begin{aligned} f_m(x) &= \int \cdots \int_{-\infty < x_1 < \cdots < x_n < \infty} n! f(x_1) \cdots f(x_n) dx_1 \cdots dx_{m-1} \\ &\quad dx_{m+1} \cdots dx_n \\ &= n! f(x) \int \cdots \int_{-\infty < x_1 < \cdots < x_{m-1} < x} f(x_1) \cdots f(x_{m-1}) dx_1 \cdots dx_{m-1} \\ &\quad \times \int \cdots \int_{x < x_{m+1} < \cdots < x_n < \infty} f(x_{m+1}) \cdots f(x_n) dx_{m+1} \cdots dx_n \end{aligned}$$

2.4.1 次序统计量的分布

$$= \frac{n!f(x)}{(m-1)!} [F(x)]^{m-1} \frac{[1-F(y)]^{n-m}}{(n-m)!}$$

=

2.4.1 次序统计量的分布

$$\begin{aligned} &= \frac{n!f(x)}{(m-1)!} [F(x)]^{m-1} \frac{[1-F(x)]^{n-m}}{(n-m)!} \\ &= m \binom{n}{m} [F(x)]^{m-1} [1-F(x)]^{n-m} f(x). \end{aligned}$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \frac{n!f(x)}{(m-1)!} [F(x)]^{m-1} \frac{[1-F(x)]^{n-m}}{(n-m)!} \\ &= m \binom{n}{m} [F(x)]^{m-1} [1-F(x)]^{n-m} f(x). \end{aligned}$$

易见其分布函数为

2.4.1 次序统计量的分布

$$\begin{aligned} &= \frac{n!f(x)}{(m-1)!} [F(x)]^{m-1} \frac{[1-F(y)]^{n-m}}{(n-m)!} \\ &= m \binom{n}{m} [F(x)]^{m-1} [1-F(x)]^{n-m} f(x). \end{aligned}$$

易见其分布函数为

$$\begin{aligned} F_m(x) &= \int_{-\infty}^x f_m(y) dy \\ &= \end{aligned}$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \frac{n!f(x)}{(m-1)!} [F(x)]^{m-1} \frac{[1-F(y)]^{n-m}}{(n-m)!} \\ &= m \binom{n}{m} [F(x)]^{m-1} [1-F(x)]^{n-m} f(x). \end{aligned}$$

易见其分布函数为

$$\begin{aligned} F_m(x) &= \int_{-\infty}^x f_m(y) dy \\ &= m \binom{n}{m} \int_{-\infty}^x [F(y)]^{m-1} [1-F(y)]^{n-m} f(y) dy \\ &= \end{aligned}$$

2.4.1 次序统计量的分布

$$\begin{aligned} &= \frac{n!f(x)}{(m-1)!} [F(x)]^{m-1} \frac{[1-F(y)]^{n-m}}{(n-m)!} \\ &= m \binom{n}{m} [F(x)]^{m-1} [1-F(x)]^{n-m} f(x). \end{aligned}$$

易见其分布函数为

$$\begin{aligned} F_m(x) &= \int_{-\infty}^x f_m(y) dy \\ &= m \binom{n}{m} \int_{-\infty}^x [F(y)]^{m-1} [1-F(y)]^{n-m} f(y) dy \\ &= m \binom{n}{m} \int_0^{F(x)} t^{m-1} (1-t)^{n-m} dt, \end{aligned}$$

2.4.1 次序统计量的分布

上式最后一个等式通过作变换 $t = F(y)$ 获得。

2.4.1 次序统计量的分布

上式最后一个等式通过作变换 $t = F(y)$ 获得。

样本极小值 $X_{(1)}$ 和极大值 $X_{(n)}$ 的密度函数可以分别将 $m = 1$ 和 $m = n$ 带入式子 (2) 得到。 □

2.4.2 极差的分布

样本极差的定义在第一章已经给出，即 $R = X_{(n)} - X_{(1)}$ 称为**样本极差**。

2.4.2 极差的分布

样本极差的定义在第一章已经给出，即 $R = X_{(n)} - X_{(1)}$ 称为**样本极差**。

为求样本极差的分布，先求 $V = X_{(j)} - X_{(i)}$ ， $j > i$ 的分布。

2.4.2 极差的分布

样本极差的定义在第一章已经给出，即 $R = X_{(n)} - X_{(1)}$ 称为**样本极差**。

为求样本极差的分布，先求 $V = X_{(j)} - X_{(i)}$ ， $j > i$ 的分布。

当取 $j = n$ ， $i = 1$ 时即可得到样本极差的分布。

2.4.2 极差的分布

具体方法如下。作下列变换：

$$\begin{cases} Z = X_{(i)}, \\ V = X_{(j)} - X_{(i)} \end{cases} \iff \begin{cases} X_{(i)} = Z, \\ X_{(j)} = Z + V. \end{cases}$$

2.4.2 极差的分布

具体方法如下。作下列变换：

$$\begin{cases} Z = X_{(i)}, \\ V = X_{(j)} - X_{(i)} \end{cases} \iff \begin{cases} X_{(i)} = Z, \\ X_{(j)} = Z + V. \end{cases}$$

变换的Jacobi行列式为 $J = \left| \frac{\partial(X_{(i)}, X_{(j)})}{\partial Z \partial V} \right| = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1$ 。

2.4.2 极差的分布

由定理2.4.2, 可知 $(X_{(i)}, X_{(j)})$ 的联合分布密度:

$$f_{ij}(x, y) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} \\ \quad \times [1 - F(y)]^{n-j} f(x) f(y), & x < y, \\ 0, & \text{其他.} \end{cases}$$

2.4.2 极差的分布

由定理2.4.2, 可知 $(X_{(i)}, X_{(j)})$ 的联合分布密度:

$$f_{ij}(x, y) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} \\ \quad \times [1 - F(y)]^{n-j} f(x)f(y), & x < y, \\ 0, & \text{其他.} \end{cases}$$

将其中 x 用 z 代替, y 用 $z + v$ 代替, 再乘上变换的Jacobi行列式的绝对值 $|J|$, 得到 (Z, V) 的联合密度

2.4.2 极差的分布

由定理2.4.2, 可知 $(X_{(i)}, X_{(j)})$ 的联合分布密度:

$$f_{ij}(x, y) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(x)]^{i-1} [F(y) - F(x)]^{j-i-1} \\ \quad \times [1 - F(y)]^{n-j} f(x) f(y), & x < y, \\ 0, & \text{其他.} \end{cases}$$

将其中 x 用 z 代替, y 用 $z + v$ 代替, 再乘上变换的Jacobi行列式的绝对值 $|J|$, 得到 (Z, V) 的联合密度

$$g_{ij}(z, v) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} [F(z)]^{i-1} [F(z+v) - F(z)]^{j-i-1} \\ \quad \times [1 - F(z+v)]^{n-j} f(z) f(z+v), v > 0, \\ 0, \text{ 其他.} \end{cases}$$

2.4.2 极差的分布

从而易知 V 的密度为

2.4.2 极差的分布

从而易知 V 的密度为

$$g_V(v) = \int_{-\infty}^{\infty} g_{ij}(z, v) dz.$$

2.4.2 极差的分布

从而易知 V 的密度为

$$g_V(v) = \int_{-\infty}^{\infty} g_{ij}(z, v) dz.$$

特别地，在 $g_{ij}(x, y)$ 的表达式中取 $i = 1, j = n$ 得到 $(Z, R) = (Z, V)$ 的联合密度为

2.4.2 极差的分布

从而易知 V 的密度为

$$g_V(v) = \int_{-\infty}^{\infty} g_{ij}(z, v) dz.$$

特别地，在 $g_{ij}(x, y)$ 的表达式中取 $i = 1, j = n$ 得到 $(Z, R) = (Z, V)$ 的联合密度为

$$g_{1n}(z, r) = \begin{cases} n(n-1)[F(z+r) - F(z)]^{n-2} f(z+r)f(z), & r > 0 \\ 0, & r \leq 0. \end{cases}$$

2.4.2 极差的分布

从而易知 V 的密度为

$$g_V(v) = \int_{-\infty}^{\infty} g_{ij}(z, v) dz.$$

特别地，在 $g_{ij}(x, y)$ 的表达式中取 $i = 1, j = n$ 得到 $(Z, R) = (Z, V)$ 的联合密度为

$$g_{1n}(z, r) = \begin{cases} n(n-1)[F(z+r) - F(z)]^{n-2} f(z+r)f(z), & r > 0 \\ 0, & r \leq 0. \end{cases}$$

而样本极差 $R = X_{(n)} - X_{(1)}$ 的边缘密度为 $g_R(r) = \int_{-\infty}^{\infty} g_{1n}(z, r) dz$ 。

2.4.3 均匀分布的情形

设 X_1, X_2, \dots, X_n , *i.i.d.* \sim 均匀分布 $U(0, 1)$, 其分布函数和密度函数分别为

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases} \quad \text{和} \quad f(x) = \begin{cases} 1, & 0 \leq x < 1, \\ 0, & \text{其他.} \end{cases}$$

2.4.3 均匀分布的情形

设 X_1, X_2, \dots, X_n , *i.i.d.* \sim 均匀分布 $U(0, 1)$, 其分布函数和密度函数分别为

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases} \quad \text{和} \quad f(x) = \begin{cases} 1, & 0 \leq x < 1, \\ 0, & \text{其他.} \end{cases}$$

设 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 X_1, X_2, \dots, X_n 的次序统计量, 由定理2.4.3中式子(2)可知单个次序统计量 $X_{(m)}$ 的密度函数为

2.4.3 均匀分布的情形

设 X_1, X_2, \dots, X_n , *i.i.d.* \sim 均匀分布 $U(0, 1)$, 其分布函数和密度函数分别为

$$F(x) = \begin{cases} 0, & x < 0, \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1. \end{cases} \quad \text{和} \quad f(x) = \begin{cases} 1, & 0 \leq x < 1, \\ 0, & \text{其他.} \end{cases}$$

设 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为样本 X_1, X_2, \dots, X_n 的次序统计量, 由定理 2.4.3 中式子 (2) 可知单个次序统计量 $X_{(m)}$ 的密度函数为

$$f_m(x) = m \binom{n}{m} x^{m-1} (1-x)^{n-m} I_{(0,1)}(x).$$

2.4.3 均匀分布的情形

由定理2.4.2中式子 (1) 可知两个 $(X_{(i)}, X_{(j)})$ ($i < j$) 的联合密度函数为

2.4.3 均匀分布的情形

由定理2.4.2中式子 (1) 可知两个 $(X_{(i)}, X_{(j)})$ ($i < j$) 的联合密度函数为

$$f_{ij}(x, y) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j}, & 0 < x < y < 1, \\ 0, & \text{其他.} \end{cases}$$

2.4.3 均匀分布的情形

由定理2.4.2中式子(1)可知两个 $(X_{(i)}, X_{(j)})$ ($i < j$) 的联合密度函数为

$$f_{ij}(x, y) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} x^{i-1} (y-x)^{j-i-1} (1-y)^{n-j}, & 0 < x < y < 1, \\ 0, & \text{其他.} \end{cases}$$

由定理2.4.1也可得到 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 的联合密度为

$$f(x_{(1)}, x_{(2)}, \dots, x_{(n)}) = \begin{cases} n!, & 0 < x_{(1)} < x_{(2)} < \dots < x_{(n)} < 1, \\ 0, & \text{其他.} \end{cases}$$

2.4.3 均匀分布的情形

同样的，还可以得到均匀分布(0, 1)情形的($Z = X_{(i)}$, $V = X_{(j)} - X_{(i)}$)的联合密度

$$g_{ij}(z, v) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} z^{i-1} v^{j-i-1} \times [1 - (z + v)]^{n-j}, & 0 < z < 1, 0 < z + v < 1, \\ 0, & \text{其他.} \end{cases}$$

2.4.3 均匀分布的情形

同样的，还可以得到均匀分布(0, 1)情形的($Z = X_{(i)}, V = X_{(j)} - X_{(i)}$)的联合密度

$$g_{ij}(z, v) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} z^{i-1} v^{j-i-1} \times [1 - (z + v)]^{n-j}, & 0 < z < 1, 0 < z + v < 1, \\ 0, & \text{其他.} \end{cases}$$

通过计算积分 $\int_0^{1-v} g_{ij}(v, z) dz$ 得 $V = X_{(j)} - X_{(i)}$ 的边缘密度为

2.4.3 均匀分布的情形

同样的，还可以得到均匀分布(0, 1)情形的($Z = X_{(i)}$, $V = X_{(j)} - X_{(i)}$)的联合密度

$$g_{ij}(z, v) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!} z^{i-1} v^{j-i-1} \times [1 - (z + v)]^{n-j}, & 0 < z < 1, 0 < z + v < 1, \\ 0, & \text{其他.} \end{cases}$$

通过计算积分 $\int_0^{1-v} g_{ij}(v, z) dz$ 得 $V = X_{(j)} - X_{(i)}$ 的边缘密度为

$$g_{ij}(v) = \begin{cases} \frac{n!}{(j-i-1)!(n-j+i)!} v^{j-i-1} (1-v)^{n-j+i}, & 0 < v < 1, \\ 0, & \text{其他.} \end{cases}$$

2.4.3 均匀分布的情形

特别地，极差 $R = X_{(n)} - X_{(1)}$ 的密度函数 $g_R(r)$ 为将上式中的 v 换成 r ，将 j 和 i 分别用 n 和 1 代替得到

2.4.3 均匀分布的情形

特别地，极差 $R = X_{(n)} - X_{(1)}$ 的密度函数 $g_R(r)$ 为将上式中的 v 换成 r ，将 j 和 i 分别用 n 和 1 代替得到

$$g_R(r) = \begin{cases} n(n-1)r^{n-2}(1-r), & 0 < r < 1, \\ 0, & \text{其他.} \end{cases}$$