数理统计
第一章
绪论

2026 年 3 月 11 日

目录

1.3.1 统计量的定义

定义 (1.3.1（统计量）)

由样本算出的量称为统计量（statistics）。或者说，统计量是样本的函数。

1.3.1 统计量的定义

定义 (1.3.1（统计量）)

由样本算出的量称为统计量（statistics）。或者说，统计量是样本的函数。

（1）统计量只与样本有关，不能与未知参数有关。

定义 (1.3.1（统计量）)

由样本算出的量称为**统计量**（statistics）。或者说，统计量是样本的函数。

（1）统计量只与样本有关，不能与未知参数有关。

例如，$X \sim N(a,\sigma^2)$，$X_1,...,X_n$是从总体X中抽取的$i.i.d.$样本，则$\sum_{i=1}^n X_i$和$\sum_{i=1}^n X_i^2$都是统计量，当a和σ^2皆为未知参数时，$\sum_{i=1}^n (X_i - a)$和$\sum_{i=1}^n \frac{X_i^2}{\sigma^2}$都不是统计量。

（2）由于样本具有两重性，它既可以看成具体的数，又可以看成随机变量（或随机向量）；统计量是样本的函数，因此统计量也具有两重性。正因为统计量可视为随机变量（或随机向量），才有概率分布可言，这是利用统计量进行统计推断的依据。

（2）由于样本具有两重性，它既可以看成具体的数，又可以看成随机变量（或随机向量）；统计量是样本的函数，因此统计量也具有两重性。正因为统计量可视为随机变量（或随机向量），才有概率分布可言，这是利用统计量进行统计推断的依据。

（3）在什么问题中选用什么统计量，要看问题的性质。一般来说，所提出的统计量应是最好地集中了样本中与所讨论问题有关的信息，这不是容易做到的。通常是从直观或一般性准则出发提出统计量，再考察它是否再某种意义下较好地集中了样本中所讨论问题有关的信息。

1. 样本均值

1. 样本均值

设$X_1, \ldots, X_n$是从某总体X中抽取的样本，则称

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

为样本均值（sample mean）。它反映了总体均值的信息。

2. 样本方差

2. 样本方差

设$X_1, \dots, X_n$是从某总体X中抽取的样本，则称

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

为样本方差（sample variance）。它反映了总体方差的信息，而S称为样本标准差，它反映了总体标准差的信息。

一些教科书上也采用

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

作为样本方差的定义。

一些教科书上也采用

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

作为样本方差的定义。

用S^2定义样本方差的好处是$E(S^2) = \sigma^2 = D(X)$（参看例3.1.1），其中$n-1$称为其自由度。

1.3.2 若干常用的统计量

样本均值和样本方差是两个最常用的统计量，它们具有如下三个性质：

样本均值和样本方差是两个最常用的统计量，它们具有如下三个性质：

（1）$\sum_{i=1}^{n}(X_i - \bar{X}) = 0.$

样本均值和样本方差是两个最常用的统计量，它们具有如下三个性质：

（1）$\sum_{i=1}^{n}(X_i - \bar{X}) = 0.$

（2）设非零实数a和b为常数，作变换$Y_i = aX_i + b$，$i = 1, 2, \ldots, n$，则$Y_1, \ldots, Y_n$的样本均值$\bar{Y} = a\bar{X} + b$，其样本方差$S_Y^2 = a^2 S_X^2$，其中S_X^2和S_Y^2分别表示$X_1, \ldots, X_n$和$Y_1, \ldots, Y_n$的样本方差。

样本均值和样本方差是两个最常用的统计量，它们具有如下三个性质：

（1）$\sum_{i=1}^{n}(X_i-\bar{X})=0.$

（2）设非零实数a和b为常数，作变换$Y_i=aX_i+b$，$i=1,2,\ldots,n$，则$Y_1,\ldots,Y_n$的样本均值$\bar{Y}=a\bar{X}+b$，其样本方差$S_Y^2=a^2S_X^2$，其中S_X^2和S_Y^2分别表示$X_1,\ldots,X_n$和$Y_1,\ldots,Y_n$的样本方差。

这个性质表明：改变度量的原点与单位，上述公式给出了新旧度量系统中样本均值和样本方差之间的关系。利用这个公式可以简化样本方差的计算。

（3）对于任何常数c，有

$$\sum_{i=1}^{n}(X_i - c)^2 \geq \sum_{i=1}^{n}(X_i - \bar{X})^2,$$

（3）对于任何常数c，有

$$\sum_{i=1}^{n}(X_i - c)^2 \geq \sum_{i=1}^{n}(X_i - \bar{X})^2,$$

且等号只在$c = \bar{X}$时成立。

（3）对于任何常数c，有

$$\sum_{i=1}^{n}(X_i-c)^2 \geq \sum_{i=1}^{n}(X_i-\bar{X})^2,$$

且等号只在$c=\bar{X}$时成立。

这个性质表明，在偏差平方和最小的准则下，用总体均值a的n次测量值的算数平均值估计a是最好的。

3. 样本矩

3. 样本矩

设$X_1, \ldots, X_n$为从总体F中抽取的样本，则称

$$a_{n,k} = \frac{1}{n} \sum_{i=1}^{n} X_i^k, \; k = 1, 2, \ldots$$

为样本k阶原点矩。

3. 样本矩

设$X_1,\ldots,X_n$为从总体F中抽取的样本，则称

$$a_{n,k}=\frac{1}{n}\sum_{i=1}^{n}X_i^k,\ k=1,2,\ldots$$

为样本k阶原点矩。

特别地，当$k=1$时，$a_{n,1}=\bar{X}$，即样本均值。

称

$$m_{n,k}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k,\ k=2,3,\ldots$$

为样本k阶中心距。

称

$$m_{n,k}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k,\ k=2,3,\ldots$$

为**样本k阶中心距**。

特别地，当$k=2$时，$m_{n,2}=\frac{(n-1)S^2}{n}$。

称

$$m_{n,k} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k,\ k = 2, 3, \ldots$$

为样本k阶中心距。

特别地，当$k = 2$时，$m_{n,2} = \frac{(n-1)S^2}{n}$。

样本的原点矩和中心矩统称为样本矩（sample moments）。

4. 二维随机向量的样本矩

4. 二维随机向量的样本矩

设$(X_1, Y_1), \ldots, (X_n, Y_n)$为从二维总体$F(x, y)$中抽取的样本，则

4. 二维随机向量的样本矩

设$(X_1, Y_1), \ldots, (X_n, Y_n)$为从二维总体$F(x, y)$中抽取的样本，则

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

4. 二维随机向量的样本矩

设$(X_1, Y_1), \ldots, (X_n, Y_n)$为从二维总体$F(x, y)$中抽取的样本，则

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,\ S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2,$$
$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,\ S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2,$$

4. 二维随机向量的样本矩

设$(X_1, Y_1), \ldots, (X_n, Y_n)$为从二维总体$F(x, y)$中抽取的样本，则

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

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,\ S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2,$$

$$S_{XY} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

4. 二维随机向量的样本矩

设$(X_1, Y_1), \ldots, (X_n, Y_n)$为从二维总体$F(x, y)$中抽取的样本，则

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

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i,\ S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2,$$

$$S_{XY} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

分别称为X和Y的样本均值、样本方差及X和Y的样本协方差（sample covariance）。

称

$$r = \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) / \sqrt{[(n-1)S_X^2] \times [(n-1)S_Y^2]}$$

为样本相关系数（sample correlation coefficient）。

5. 次序统计量及其有关统计量

5. 次序统计量及其有关统计量

设$X_1, \ldots, X_n$为从总体F中抽取的样本，将其按大小排列为$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$，则$(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$称为样本$(X_1, \ldots, X_n)$的次序统计量（order statistics）。

5. 次序统计量及其有关统计量

设$X_1, \ldots, X_n$为从总体F中抽取的样本，将其按大小排列为$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$，则$(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$称为样本$(X_1, \ldots, X_n)$的次序统计量（order statistics）。

$(X_{(1)}, \ldots, X_{(n)})$的任一部分也称为次序统计量。

（1）样本中位数：

（1）样本中位数：

$$m_{\frac{1}{2}} = \begin{cases} X_{((n+1)/2)}, & n\text{为奇数} \\ \frac{1}{2}[X_{(n/2)} + X_{(n/2+1)}], & n\text{为偶数} \end{cases}$$

（1）样本中位数：

$$m_{\frac{1}{2}} = \begin{cases} X_{((n+1)/2)}, & n\text{为奇数} \\ \frac{1}{2}[X_{(n/2)} + X_{(n/2+1)}], & n\text{为偶数} \end{cases}$$

称为样本中位数（sample median），它反映总体中位数的信息。

（1）样本中位数：

$$m_{\frac{1}{2}}=\begin{cases}X_{((n+1)/2)}, & n\text{为奇数}\\ \frac{1}{2}[X_{(n/2)}+X_{(n/2+1)}], & n\text{为偶数}\end{cases}$$

称为样本中位数（sample median），它反映总体中位数的信息。

当总体分布关于某点对称时，对称中心既是总体中位数，又是总体均值。此时$m_{1/2}$也反映总体均值的信息。

（2）**极值**：$X_{(1)}$和$X_{(n)}$称为样本的极小值和极大值，它们统称为样本极值（extremum of sample）。

（2）**极值**：$X_{(1)}$和$X_{(n)}$称为样本的极小值和极大值，它们统称为样本极值（extremum of sample）。

极值统计量在关于灾害问题和材料试验的统计分析中是常用的统计量。

（3）样本p分位数（$0 < p < 1$）可定义为

（3）样本p分位数（$0 < p < 1$）可定义为

$$X_{(m)} = \begin{cases} X_{[np+1]}, & np\text{为非整数} \\ \frac{1}{2}(X_{[np]} + X_{[np+1]}), & np\text{为整数}. \end{cases}$$

（3）**样本p分位数**（$0 < p < 1$）可定义为

$$X_{(m)} = \begin{cases} X_{[np+1]}, & np\text{为非整数} \\ \frac{1}{2}(X_{[np]} + X_{[np+1]}), & np\text{为整数}. \end{cases}$$

此处$[a]$表示实数a的整数部分。

（3）样本p分位数（$0 < p < 1$）可定义为

$$X_{(m)} = \begin{cases} X_{[np+1]}, & np\text{为非整数} \\ \frac{1}{2}(X_{[np]} + X_{[np+1]}), & np\text{为整数}. \end{cases}$$

此处$[a]$表示实数a的整数部分。

样本p分位数（sample p-fractile）反映了总体p分位数信息。

样本的p分位数给定的方式也有多种。

样本的p分位数给定的方式也有多种。

若设x_p为分布F（密度函数为f）的下侧p分位数，则需满足

样本的p分位数给定的方式也有多种。

若设x_p为分布F（密度函数为f）的下侧p分位数，则需满足

$$F(x_p) = \int_{-\infty}^{x_p} f(x)dx = p,\ p \in (0,1).$$

样本的p分位数给定的方式也有多种。

若设x_p为分布F（密度函数为f）的下侧p分位数，则需满足

$$F(x_p) = \int_{-\infty}^{x_p} f(x)dx = p,\ p \in (0,1).$$

若设x'_p为分布F（密度函数为f）的上侧p分位数，则需满足

样本的p分位数给定的方式也有多种。

若设x_p为分布F（密度函数为f）的下侧p分位数，则需满足

$$F(x_p)=\int_{-\infty}^{x_p}f(x)dx=p,\ p\in(0,1).$$

若设x'_p为分布F（密度函数为f）的上侧p分位数，则需满足

$$1-F(x'_p)=\int_{x'_p}^{\infty}f(x)dx=p,\ p\in(0,1).$$

还有一种比较常见的样本的p分位数的定义是

还有一种比较常见的样本的p分位数的定义是

$$x_p = \inf\{x : F_n(x) \geq p\}.$$

还有一种比较常见的样本的p分位数的定义是

$$x_p = \inf\{x : F_n(x) \geq p\}.$$

这里$F_n(x) = \frac{1}{n} I_{\{X_i \leq x\}}$是样本的经验分布函数。

（4）样本极差：$R = X_{(n)} - X_{(1)}$称为**样本极差**（sample range），它是反映总体分布散布程度的信息。

方差（或标准差）反映了随机变量取值的波动程度，但在比较两个随机变量的波动大小时，如果仅看方差（或标准差）的大小有时会产生不合理的现象。

方差（或标准差）反映了随机变量取值的波动程度，但在比较两个随机变量的波动大小时，如果仅看方差（或标准差）的大小有时会产生不合理的现象。

这里有两个原因：（1）随机变量的取值有量纲，不同量纲的随机变量用其方差（标准差）去比较它们的波动大小不太合理。

方差（或标准差）反映了随机变量取值的波动程度，但在比较两个随机变量的波动大小时，如果仅看方差（或标准差）的大小有时会产生不合理的现象。

这里有两个原因：（1）随机变量的取值有量纲，不同量纲的随机变量用其方差（标准差）去比较它们的波动大小不太合理。

（2）在取值的量纲相同的情况下，取值的大小有一个相对性的问题，取值较大的随机变量的方差（标准差）也允许大一些。

6. 样本变异系数

6. 样本变异系数

设$X_1, \ldots, X_n$为从总体F中抽取的样本，则称

$$\hat{\nu} = \frac{S_n}{\bar{X}}.$$

为样本变异系数（sample coefficient of variation）。

6. 样本变异系数

设$X_1, \ldots, X_n$为从总体F中抽取的样本，则称

$$\hat{\nu} = \frac{S_n}{\bar{X}}.$$

为样本变异系数（sample coefficient of variation）。

它反映了总体变异系数的信息。

6. 样本变异系数

设$X_1, \ldots, X_n$为从总体F中抽取的样本，则称

$$\hat{\nu} = \frac{S_n}{\bar{X}}.$$

为样本变异系数（sample coefficient of variation）。

它反映了总体变异系数的信息。

总体变异系数的定义是$\nu = \frac{\sqrt{Var(X)}}{E(X)}$，它是衡量总体分布散布程度的量，但这散布程度是以总体均值为单位来度量。

因为变异系数的分母是以其数学期望为单位去度量随机变量取值波动程度的特征数，而标准差的量纲与数学期望的量纲是一致的，所以变异系数是一个无量纲的量，从而消除量纲对波动的影响。

假设比较X和Y的波动，已知X和Y的单位是相同的，且

假设比较X和Y的波动，已知X和Y的单位是相同的，且

$$E(X) = 10,\ \ Var(X) = 1$$

$$E(Y) = 1,\ \ Var(Y) = 0.04$$

假设比较X和Y的波动，已知X和Y的单位是相同的，且

$$E(X) = 10,\ \ Var(X) = 1$$

$$E(Y) = 1,\ \ Var(Y) = 0.04$$

比较X和Y的方差，可以认为Y的波动小吗?

假设比较X和Y的波动，已知X和Y的单位是相同的，且

$$E(X) = 10, \ \ Var(X) = 1$$

$$E(Y) = 1, \ \ Var(Y) = 0.04$$

比较X和Y的方差，可以认为Y的波动小吗？

- 若直接比较方差大小，$Var(X) > Var(Y)$，认为X波动大。

假设比较X和Y的波动，已知X和Y的单位是相同的，且

$$E(X) = 10,\ \ Var(X) = 1$$

$$E(Y) = 1,\ \ Var(Y) = 0.04$$

比较X和Y的方差，可以认为Y的波动小吗？

- 若直接比较方差大小，$Var(X) > Var(Y)$，认为X波动大。
- 若比较变异系数，$\nu_X = \frac{\sqrt{Var(X)}}{E(X)} = \frac{1}{10} = 0.1$，$\nu_Y = \frac{\sqrt{Var(Y)}}{E(Y)} = \frac{\sqrt{0.04}}{1} = 0.2$，说明$Y$比$X$的波动程度大。 √

7. 样本偏度

7. 样本偏度

设$X_1,\dots,X_n$为从总体F中抽取的样本，则称

$$\hat{\beta}_1 = \frac{m_{n,3}}{m_{n,2}^{3/2}} = \sqrt{n}\sum_{i=1}^{n}(X_i - \bar{X})^3/(\sum_{i=1}^{n}(X_i - \bar{X}))^{3/2}$$

为样本偏度（sample skewness）。

7. 样本偏度

设$X_1, \ldots, X_n$为从总体F中抽取的样本，则称

$$\hat{\beta}_1 = \frac{m_{n,3}}{m_{n,2}^{3/2}} = \sqrt{n}\sum_{i=1}^{n}(X_i - \bar{X})^3/(\sum_{i=1}^{n}(X_i - \bar{X}))^{3/2}$$

为样本偏度（sample skewness）。 它反映了总体偏度的信息。

总体偏度的定义是$\beta_1 = \mu_3/\mu_2^{3/2}$，此处$\mu_i\ (i = 2, 3)$是总体的$i$阶中心距。

总体偏度的定义是$\beta_1 = \mu_3/\mu_2^{3/2}$，此处$\mu_i\ (i = 2, 3)$是总体的$i$阶中心距。

β_1是反映总体分布的非对称性或“偏倚性”的一种度量。

总体偏度的定义是$\beta_1 = \mu_3/\mu_2^{3/2}$，此处$\mu_i\ (i = 2, 3)$是总体的$i$阶中心距。

β_1是反映总体分布的非对称性或“偏倚性”的一种度量。

- 当$\beta_1 > 0$，总体分布右偏（正偏）；

总体偏度的定义是$\beta_1 = \mu_3/\mu_2^{3/2}$，此处$\mu_i\ (i=2,3)$是总体的$i$阶中心距。

β_1是反映总体分布的非对称性或“偏倚性”的一种度量。

- 当$\beta_1 > 0$，总体分布右偏（正偏）；
- 当$\beta_1 = 0$时，总体分布无偏（对称）；

总体偏度的定义是$\beta_1 = \mu_3/\mu_2^{3/2}$，此处$\mu_i\ (i = 2, 3)$是总体的$i$阶中心距。

β_1是反映总体分布的非对称性或“偏倚性”的一种度量。

- 当$\beta_1 > 0$，总体分布右偏（正偏）；
- 当$\beta_1 = 0$时，总体分布无偏（对称）；
- 当$\beta_1 < 0$时，总体分布左偏（负偏）。

1.3.2 若干常用的统计量

总体偏度的定义是$\beta_1 = \mu_3/\mu_2^{3/2}$，此处$\mu_i\ (i = 2, 3)$是总体的$i$阶中心距。

β_1是反映总体分布的非对称性或“偏倚性”的一种度量。

- 当$\beta_1 > 0$，总体分布右偏（正偏）；
- 当$\beta_1 = 0$时，总体分布无偏（对称）；
- 当$\beta_1 < 0$时，总体分布左偏（负偏）。

正态分布$N(a, \sigma^2)$的偏度为零。

8. 样本峰度

8. 样本峰度

设$X_1, \ldots, X_n$为从总体F中抽取的样本，则称

$$\hat{\beta}_2 = \frac{m_{n,4}}{m_{n,2}^2} - 3 = n\sum_{i=1}^{n}(X_i - \bar{X})^4/(\sum_{i=1}^{n}(X_i - \bar{X})^2)^2 - 3$$

为样本峰度（sample kurtosis）。

8. 样本峰度

设$X_1, \ldots, X_n$为从总体F中抽取的样本，则称

$$\hat{\beta}_2 = \frac{m_{n,4}}{m_{n,2}^2} - 3 = n\sum_{i=1}^{n}(X_i - \bar{X})^4 / (\sum_{i=1}^{n}(X_i - \bar{X})^2)^2 - 3$$

为样本峰度（sample kurtosis）。 它反映了总体峰度的信息。

1.3.2 若干常用的统计量

总体峰度的定义是$\beta_2 = \mu_4/\mu_2^2 - 3$，此处$\mu_i$ $(i = 2, 4)$是总体的i阶中心距。

β_2是反映总体分布的密度函数在众数（即密度函数的最大值点）附件"峰"的尖峭程度的一种度量。

总体峰度的定义是$\beta_2 = \mu_4/\mu_2^2 - 3$，此处$\mu_i\ (i = 2, 4)$是总体的$i$阶中心距。

β_2是反映总体分布的密度函数在众数（即密度函数的最大值点）附件"峰"的尖峭程度的一种度量。

- 当$\beta_2 > 0$，总体分布更尖峭，尾部更粗（t分布）;

总体峰度的定义是$\beta_2 = \mu_4/\mu_2^2 - 3$，此处$\mu_i \ (i = 2, 4)$是总体的$i$阶中心距。

β_2是反映总体分布的密度函数在众数（即密度函数的最大值点）附件"峰"的尖峭程度的一种度量。

- 当$\beta_2 > 0$，总体分布更尖峭，尾部更粗（t分布）；
- 当$\beta_2 = 0$时，总体分布与正态相当；

总体峰度的定义是$\beta_2 = \mu_4/\mu_2^2 - 3$，此处$\mu_i \ (i = 2, 4)$是总体的$i$阶中心距。

β_2是反映总体分布的密度函数在众数（即密度函数的最大值点）附件"峰"的尖峭程度的一种度量。

- 当$\beta_2 > 0$，总体分布更尖峭，尾部更粗（t分布）；
- 当$\beta_2 = 0$时，总体分布与正态相当；
- 当$\beta_2 < 0$时，总体分布更平坦，尾部更细（均匀分布）。

总体峰度的定义是$\beta_2 = \mu_4/\mu_2^2 - 3$，此处$\mu_i\ (i=2,4)$是总体的$i$阶中心距。

β_2是反映总体分布的密度函数在众数（即密度函数的最大值点）附件"峰"的尖峭程度的一种度量。

- 当$\beta_2 > 0$，总体分布更尖峭，尾部更粗（t分布）；
- 当$\beta_2 = 0$时，总体分布与正态相当；
- 当$\beta_2 < 0$时，总体分布更平坦，尾部更细（均匀分布）。

正态分布$N(a, \sigma^2)$的峰度为零。

定义 (1.3.2)

设$X_1, \ldots, X_n$为自总体$F(x)$中抽取的$i.i.d.$样本，将其按大小排列为$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$，对任意实数$x$，称下列函数：

定义 (1.3.2)

设$X_1,\dots,X_n$为自总体$F(x)$中抽取的$i.i.d.$样本，将其按大小排列为$X_{(1)}\le X_{(2)}\le\cdots\le X_{(n)}$，对任意实数$x$，称下列函数：

$$F_n(x)=\begin{cases}0,\ \ x<X_{(1)},\\ \frac{k}{n},\ \ X_{(k)}\le x<X_{(k+1)},\ \ k=1,2,\dots,n-1,\\ 1,\ \ X_{(n)}\le x\end{cases}$$

为经验分布函数。

易见经验分布函数是单调、非降、右连续函数，具有分布函数的基本性质。

易见经验分布函数是单调、非降、右连续函数，具有分布函数的基本性质。

它在$x = x_{(k)}$，$k = 1, 2, \ldots, n$处有间断，它是在每个间断点跳跃的幅度为$1/n$的阶梯函数。即

$$P(X = x_{(k)}) = \frac{1}{n},\ k = 1, \ldots, n.$$

1.3.3 经验分布函数

$F_n(x)$可以看成总体分布函数$F(x) = P(X \leq x)$的一个估计量。

1.3.3 经验分布函数

$F_n(x)$可以看成总体分布函数$F(x)=P(X\leq x)$的一个估计量。

若记示性函数

$$I_A(x)=\begin{cases}1, & x\in A\\ 0, & \text{其他},\end{cases}$$

$F_n(x)$可以看成总体分布函数$F(x)=P(X\leq x)$的一个估计量。

若记示性函数

$$I_A(x)=\begin{cases}1, & x\in A\\ 0, & \text{其他},\end{cases}$$

则$F_n(x)$可表示为

1.3.3 经验分布函数

$F_n(x)$可以看成总体分布函数$F(x) = P(X \leq x)$的一个估计量。

若记示性函数

$$I_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{其他}, \end{cases}$$

则$F_n(x)$可表示为

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(-\infty, x]}(X_i).$$

$F_n(x)$可以看成总体分布函数$F(x) = P(X \leq x)$的一个估计量。

若记示性函数

$$I_A(x) = \begin{cases} 1, & x \in A \\ 0, & \text{其他}, \end{cases}$$

则$F_n(x)$可表示为

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(-\infty, x]}(X_i).$$

由定义可知$F_n(x)$是仅依赖于样本$X_1, X_2, \ldots, X_n$的函数，因此它是统计量。

它可能取值为$0, 1/n, 2/n, \ldots, (n-1)/n, 1$。

它可能取值为$0, 1/n, 2/n, \ldots, (n-1)/n, 1$。

若记$Y_i = I_{(-\infty,x]}(X_i),\ i = 1, 2, \ldots, n$，则

$$
\begin{aligned}
P(Y_i = 1) &= P(X_i \leq x) = F(x), \\
P(Y_i = 0) &= 1 - F(x),
\end{aligned}
$$

且$Y_1, Y_2, \ldots, Y_n\ i.i.d. \sim b(1, F(x))$。

它可能取值为$0, 1/n, 2/n, \ldots, (n-1)/n, 1$。

若记$Y_i = I_{(-\infty,x]}(X_i),\ i = 1, 2, \ldots, n$，则

$$
\begin{aligned}
P(Y_i = 1) &= P(X_i \le x) = F(x), \\
P(Y_i = 0) &= 1 - F(x),
\end{aligned}
$$

且$Y_1, Y_2, \ldots, Y_n\ i.i.d. \sim b(1, F(x))$。

故$nF_n(x) = \sum_{i=1}^n I_{(-\infty,x]}(X_i) = \sum_{i=1}^n Y_i \sim b(n, F(x))$。

令$\sum_{i=1}^{n} Y_i = k$，对$k = 0, 1, \ldots, n$有

令$\sum_{i=1}^{n} Y_i = k$，对$k = 0, 1, \ldots, n$有

$$P(\sum_{i=1}^{n} Y_i = k) = P(F_n(x) = \frac{k}{n}) = \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k}.$$

利用二项分布的性质可知对任一固定的$x \in (-\infty, \infty)$,

$F_n(x) = \frac{\sum_{i=1}^{n} Y_i}{n} = \bar{Y}$具有如下的大样本性质:

1.3.3 经验分布函数

利用二项分布的性质可知对任一固定的$x \in (-\infty, \infty)$，

$F_n(x) = \frac{\sum_{i=1}^{n} Y_i}{n} = \bar{Y}$具有如下的大样本性质:

（1）由中心极限定理，则当$n \to \infty$时有

1.3.3 经验分布函数

利用二项分布的性质可知对任一固定的$x \in (-\infty, \infty)$，
$F_n(x) = \frac{\sum_{i=1}^n Y_i}{n} = \bar{Y}$具有如下的大样本性质:

（1）由中心极限定理，则当$n \to \infty$时有

$$\frac{\sqrt{n}(F_n(x) - F(x))}{\sqrt{F(x)(1 - F(x))}} \xrightarrow{\mathcal{L}} N(0, 1);$$

此处$\xrightarrow{\mathcal{L}}$表示依分布收敛。

（2）由Bernoulli（或辛钦）大数定律（《概率论》p.g.321），则在$n \to \infty$时有

（2）由Bernoulli（或辛钦）大数定律（《概率论》p.g.321），则在$n \to \infty$时有

$$F_n(x) \xrightarrow{p} F(x);$$

（2）由Bernoulli（或辛钦）大数定律（《概率论》p.g.321），则在$n \to \infty$时有

$$F_n(x) \xrightarrow{p} F(x);$$

（3）由Borel强大数定律（附录A11，p.g.87），则有

$$P(\lim_{n\to\infty} F_n(x) = F(x)) = 1;$$

（4）更进一步，有下列格里汶科定理（Glivenko-Cantelli Theorem）

定理 (1.3.1)

设$F(x)$为$r.v.\ X$的分布函数，$X_1, \ldots, X_n$为取自总体$F(x)$的简单随机样本，$F_n(x)$为其经验分布函数，记$D_n = \sup\limits_{-\infty<x<\infty} |F_n(x) - F(x)|$，则有

$$P(\lim_{n\to\infty} D_n = 0) = 1.$$

（4）更进一步，有下列格里汶科定理（Glivenko-Cantelli Theorem）

定理 (1.3.1)

设$F(x)$为$r.v.\ X$的分布函数，$X_1,\ldots,X_n$为取自总体$F(x)$的简单随机样本，$F_n(x)$为其经验分布函数，记$D_n=\sup\limits_{-\infty<x<\infty}|F_n(x)-F(x)|$，则有

$$P(\lim_{n\to\infty}D_n=0)=1.$$

注 1.3.1 上述定理中的D_n可用来衡量$F_n(x)$和$F(x)$之间在所有的x值上最大的差异程度，格里汶科定理表明：当n足够大时，对所有的x值，经验分布函数$F_n(x)$与理论分布函数$F(x)$之间只有很小的差别。

目录

样本既然是随机变量，就有一定的概率分布，这个概率分布就称为样本分布（sample distribution）。

1.4.1 统计模型的确定方法及例子

样本既然是随机变量，就有一定的概率分布，这个概率分布就称为样本分布（sample distribution）。

一个问题的统计模型（statistical model），就是指研究该问题时所抽样本的分布。

例 (1.4.1)

一大批产品共有N件，其中废品M件，N已知，而M未知。现在从中抽取n个检验其中废品的件数，用以估计M或废品率$p=M/N$。抽样方式为：不放回抽样，一次抽一个，依次抽取，直到抽完n个为止。求样本分布。

解

例 (1.4.1)

一大批产品共有N件，其中废品M件，N已知，而M未知。现在从中抽取n个检验其中废品的件数，用以估计M或废品率$p = M/N$。抽样方式为：不放回抽样，一次抽一个，依次抽取，直到抽完n个为止。求样本分布。

解 设X_i表示第i次抽出的样本，令

$$X_i = \begin{cases} 1, & \text{第i次抽出的为废品}, \\ 0, & \text{第i次抽出的为合格品}. \end{cases}$$

样本$X_1, X_2, \cdots, X_n$中的每一个都只能取$0, 1$两个值之一。

样本$X_1, X_2, \cdots, X_n$中的每一个都只能取$0, 1$两个值之一。

给定一组样本$x_1, x_2, \cdots, x_n$，每个x_i为0或1。

1.4.1 统计模型的确定方法及例子

样本$X_1, X_2, \cdots, X_n$中的每一个都只能取$0, 1$两个值之一。

给定一组样本$x_1, x_2, \cdots, x_n$，每个x_i为0或1。

所求的样本分布为$P(X_1 = x_1, \cdots, X_n = x_n)$。

样本$X_1, X_2, \cdots, X_n$中的每一个都只能取$0, 1$两个值之一。

给定一组样本$x_1, x_2, \cdots, x_n$，每个x_i为0或1。

所求的样本分布为$P(X_1 = x_1, \cdots, X_n = x_n)$。

若记事件$A_i = \{X_i = x_i\}$，利用概率乘法公式

$$P(A_1 \cdots A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 A_2 \cdots A_{n-1}),$$

不难求出样本分布。

记$\sum_{i=1}^{n} x_i = a$，利用公式可得

记$\sum_{i=1}^{n} x_i = a$，利用公式可得

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n)$$

$$=$$

记$\sum_{i=1}^{n} x_i = a$，利用公式可得

$$
\begin{aligned}
& P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) \\
= & P(X_1 = x_1) P(X_2 = x_2 | X_1 = x_1) P(X_3 = x_3 | X_2 = x_2, X_1 = x_1) \\
& \times \cdots \times P(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_1 = x_1)
\end{aligned}
$$

$$=$$

记$\sum_{i=1}^{n} x_i = a$，利用公式可得

$$
\begin{aligned}
& P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) \\
= & P(X_1 = x_1)P(X_2 = x_2 | X_1 = x_1)P(X_3 = x_3 | X_2 = x_2, X_1 = x_1) \\
& \times \cdots \times P(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_1 = x_1) \\
= & \begin{cases} \frac{(N-M)!(N-n)!}{(N-M-n)!N!}, & a = 0, \\ \frac{M}{N} \cdot \frac{M-1}{N-1} \cdots \frac{M-(a-1)}{N-(a-1)} \cdot \frac{N-M}{N-a} \cdots \frac{N-M-(n-a-1)}{N-(n-1)}, & 0 < a < n, \\ \frac{M!(N-n)!}{(M-n)!N!}, & a = n, \\ 0, & \text{其余情形}. \end{cases}
\end{aligned} \tag{1}
$$

由上述计算可见样本$X_1, X_2, \ldots, X_n$不是相互独立的，样本分布是利用乘法公式，通过条件概率计算出来的。

由上述计算可见样本$X_1, X_2, \ldots, X_n$不是相互独立的，样本分布是利用乘法公式，通过条件概率计算出来的。

例 (1.4.2)

以抽样方式改为有放回抽样，即每次抽样后记下结果，然后将其放回去，再抽第二个，直到抽完n个为止，求样本分布。

由上述计算可见样本$X_1, X_2, \ldots, X_n$不是相互独立的，样本分布是利用乘法公式，通过条件概率计算出来的。

例 (1.4.2)

以抽样方式改为有放回抽样，即每次抽样后记下结果，然后将其放回去，再抽第二个，直到抽完n个为止，求样本分布。

解 在有放回抽样情形，每次抽样时，N个产品中的每一个皆以$1/N$的概率被抽出，此时$P(X_i = 1) = M/N$，$P(X_i = 0) = (N - M)/N$，故有

$$P(X_1 = x_1, \cdots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n)$$

$$=$$

1.4.1 统计模型的确定方法及例子

$$
\begin{aligned}
&P(X_1 = x_1, \cdots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n) \\
&= \begin{cases} (\frac{M}{N})^a (\frac{M-N}{N})^{n-a}, & x_1, \ldots, x_n \text{为}0\text{或}1\text{且} \sum_{i=1}^n x_i = a, \\ 0, & \text{其余情形}. \end{cases}
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
&P(X_1=x_1,\cdots,X_n=x_n)=P(X_1=x_1)\cdots P(X_n=x_n)\\
&=\begin{cases}(\frac{M}{N})^a(\frac{M-N}{N})^{n-a}, & x_1,\ldots,x_n\text{为}0\text{或}1\text{且}\sum_{i=1}^n x_i=a,\\ 0, & \text{其余情形}.\end{cases}
\end{aligned} \tag{2}
$$

可见此例比例1.4.1要简单，因为本例中样本$X_1,\cdots,X_n$是独立同分布的，而例1.4.1中$X_1,\cdots,X_n$不独立。

$$
\begin{aligned}
&P(X_1 = x_1, \cdots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n) \\
&= \begin{cases} (\frac{M}{N})^a (\frac{M-N}{N})^{n-a}, & x_1, \ldots, x_n \text{为0或1且} \sum_{i=1}^n x_i = a, \\ 0, & \text{其余情形}. \end{cases}
\end{aligned} \tag{2}
$$

可见此例比例1.4.1要简单，因为本例中样本$X_1, \cdots, X_n$是独立同分布的，而例1.4.1中$X_1, \cdots, X_n$不独立。

当n/N很小时，式子（1）和（2）差别很小，因而当n/N很小时，可把例1.4.1中的无放回抽样当成有放回抽样来处理。

例 (1.4.3)

为估计一物件的重量a，用一架天平将它重复称n次，结果记为$X_1, X_2, \ldots, X_n$，求样本$X_1, \ldots, X_n$的联合分布。

例 (1.4.3)

为估计一物件的重量a，用一架天平将它重复称n次，结果记为$X_1, X_2, \ldots, X_n$，求样本$X_1, \ldots, X_n$的联合分布。

解

假设1：各次称重独立进行

假设2：各次称重在”相同条件“下进行

假设3：称重误差服从$N(0, \sigma^2)$，即$X_i \sim N(a, \sigma^2)$

例 (1.4.3)

为估计一物件的重量a，用一架天平将它重复称n次，结果记为$X_1, X_2, \ldots, X_n$，求样本$X_1, \ldots, X_n$的联合分布。

解

假设1：各次称重独立进行

假设2：各次称重在”相同条件“下进行

假设3：称重误差服从$N(0, \sigma^2)$，即$X_i \sim N(a, \sigma^2)$

$$f_{X_i}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{(x-a)^2}{2\sigma^2}\}.$$

假设4：天平没有系统误差

假设4：天平没有系统误差

因此简单随机样本$X_1, X_2, \ldots, X_n$的分布为

假设4：天平没有系统误差

因此简单随机样本$X_1, X_2, \ldots, X_n$的分布为

$$f(x_1, \cdots, x_n) = (\sqrt{2\pi}\sigma)^{-n} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - a)^2\}.$$

例 (1.4.4)

某工厂生产一种电子元件，如晶体管。由于大批量生产受随机因素的干扰，生产出的晶体管寿命不同。从中抽取n个做寿命实验，用以估计其平均寿命，求样本的分布。

例 (1.4.4)

某工厂生产一种电子元件，如晶体管。由于大批量生产受随机因素的干扰，生产出的晶体管寿命不同。从中抽取n个做寿命实验，用以估计其平均寿命，求样本的分布。

解

将其在正常条件下使用，直到n个元件全部失效为止，测得n个元件的寿命为$X_1, \ldots, X_n$，这就是样本。

例 (1.4.4)

某工厂生产一种电子元件，如晶体管。由于大批量生产受随机因素的干扰，生产出的晶体管寿命不同。从中抽取n个做寿命实验，用以估计其平均寿命，求样本的分布。

解

将其在正常条件下使用，直到n个元件全部失效为止，测得n个元件的寿命为$X_1, \ldots, X_n$，这就是样本。

假设1：工厂的生产量大，生产条件在所考察的那段时间是稳定的。可以认为$X_1, \cdots, X_n$是相互独立相同分布的随机变量；

假设2：假定寿命分布是无记忆性，即在元件已经使用了长为t的一段时间尚未失效的条件下，元件至少还能使用一段时间s的概率与t无关。

1.4.1 统计模型的确定方法及例子

假设2：假定寿命分布是无记忆性，即在元件已经使用了长为t的一段时间尚未失效的条件下，元件至少还能使用一段时间s的概率与t无关。

在概率论（课本p.g. 138，课件12）中已经证明了：在上述两个假定下

假设2：假定寿命分布是无记忆性，即在元件已经使用了长为t的一段时间尚未失效的条件下，元件至少还能使用一段时间s的概率与t无关。

在概率论（课本p.g. 138，课件12）中已经证明了：在上述两个假定下

$$P(X_1 < x) = \begin{cases} 1 - e^{-\lambda x}. \ x \geq 0, \\ 0, \ x \leq 0, \end{cases}$$

即X_1服从指数分布。

也可以写成概率密度的形式：X_1有概率密度

$$f(x,\lambda)=\begin{cases}\lambda e^{-\lambda x},\ x>0;\\ 0,\ x\le 0.\end{cases}$$

也可以写成概率密度的形式：X_1有概率密度

$$f(x,\lambda)=\begin{cases}\lambda e^{-\lambda x},\ x>0;\\ 0,\ x\le 0.\end{cases}$$

因此，样本$X_1, X_2, \ldots, X_n$的联合密度为

$$f(x_1,\ldots,x_n;\lambda)=\begin{cases}\lambda^n e^{-n\lambda\bar{x}},\ x_1,x_2,\ldots,x_n>0;\\ 0,\ \text{其他}.\end{cases}$$

此处，$\bar{x}=\frac{\sum_{i=1}^n x_i}{n}$。

例 (1.4.5)

若样本是多维的，以上讨论也适用。例如，在一群人中抽取n个，测其身高和体重得$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$，这就是大小为$n$的二维样本。在一定条件下，他们是$i.i.d.$的，且$(X_1, Y_1)$的分布为二维正态分布$N(a, b; \sigma_1^2, \sigma_2^2; \rho)$，记$\theta = (a, b, \sigma_1^2, \sigma_2^2, \rho)$，则样本$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$的联合密度是什么？

解

由二维正态分布的概率密度函数和题目中独立同分布的假设，可知样本的联合密度函数为

解

由二维正态分布的概率密度函数和题目中独立同分布的假设，可知样本的联合密度函数为

$$f(x_1,y_1,\ldots,x_n,y_n;\theta)=\left(\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\right)^n$$
$$\times\prod_{i=1}^{n}\{\exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x_i-a)^2}{\sigma_1^2}-\frac{2\rho(x_i-a)(y_i-b)}{\sigma_1\sigma_2}+\frac{(y_i-b)^2}{\sigma_2^2}]\}\}.$$

1.4.2 有关统计模型的几点说明

1.统计模型的名称

1.统计模型的名称

一个问题的**统计模型**（statistical model），就是指研究该问题时所抽样本的分布，也常称为概率模型。

1.4.2 有关统计模型的几点说明

1.统计模型的名称

一个问题的**统计模型**（statistical model），就是指研究该问题时所抽样本的分布，也常称为概率模型。

由于模型只取决于样本的分布，故常把分布的名称作为模型的名称。

1.统计模型的名称

一个问题的**统计模型**（statistical model），就是指研究该问题时所抽样本的分布，也常称为概率模型。

由于模型只取决于样本的分布，故常把分布的名称作为模型的名称。

例如，正态分布模型简称为正态模型，Poisson分布模型简称为Poisson模型，指数分布模型简称为指数模型等。

2."问题的模型"与抽样方式

2." 问题的模型" 与抽样方式

统计分析基于样本，确定统计问题需要先确定样本的产生方法，即确定样本的分布。

2." 问题的模型" 与抽样方式

统计分析基于样本，确定统计问题需要先确定样本的产生方法，即确定样本的分布。

如例1.4.1和例1.4.2中采取有放回和不放回两种不同的抽样方式，就确定了两种不同的样本的分布，即统计模型。

3.很多性质不一样的问题，可以归入到同一模型下

1.4.2 有关统计模型的几点说明

3.很多性质不一样的问题，可以归入到同一模型下

数理统计学的任务就是研究种种统计模型中所能提出来的种种统计问题。

3.很多性质不一样的问题，可以归入到同一模型下

数理统计学的任务就是研究种种统计模型中所能提出来的种种统计问题。

许多不同专业部门中性质不一样的问题在一定条件下可归入到同一模型下，对这一模型种种统计推断问题的研究反过来又可应用于解决这些专业部门提出的性质各异的同一类问题。

4.同一模型下可以提出很多不同的统计推断问题

4.同一模型下可以提出很多不同的统计推断问题

例如在例子1.4.3中，问题是估计重物a，为了考察天平的精度，可以提出估计σ^2，还可以对a和σ^2提出假设检验和区间估计的问题等。

目录

在统计理论问题中，许多统计推断方法的优良性，对一类范围广泛的统计模型（亦称为分布族）有较满意的结果，这类分布族为指数型分布族（简称指数族）。

在统计理论问题中，许多统计推断方法的优良性，对一类范围广泛的统计模型（亦称为分布族）有较满意的结果，这类分布族为指数型分布族（简称指数族）。

常见的分布，如正态分布、二项分布、Poisson分布、负二项分布、指数分布、Gamma分布和多项分布等都属于这一分布族。

在统计理论问题中，许多统计推断方法的优良性，对一类范围广泛的统计模型（亦称为分布族）有较满意的结果，这类分布族为指数型分布族（简称指数族）。

常见的分布，如正态分布、二项分布、Poisson分布、负二项分布、指数分布、Gamma分布和多项分布等都属于这一分布族。

这些表面上看起来各不相同的分布族，其实都可以统一在一种包罗更广的一类称为指数型分布族的模式中。对这一类分布族的讨论，可以抓住他们的共性，并且许多统计理论问题，对指数族都有较彻底的解决。

定义 (1.5.1)

设$\mathcal{F} = \{f(x;\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$是定义在样本空间$\mathcal{X}$上的分布族，其中$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$，$\Theta$为参数空间。若其概率函数$f(x;\boldsymbol{\theta})$可表示成如下形式：

定义 (1.5.1)

设$\mathcal{F} = \{f(x;\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$是定义在样本空间$\mathcal{X}$上的分布族，其中$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$，$\Theta$为参数空间。若其概率函数$f(x;\boldsymbol{\theta})$可表示成如下形式：

$$f(x;\boldsymbol{\theta}) = C(\boldsymbol{\theta}) \exp\{\sum_{i=1}^{k} Q_i(\boldsymbol{\theta}) T_i(x)\} h(x),$$

定义 (1.5.1)

设$\mathcal{F}=\{f(x;\boldsymbol{\theta}):\boldsymbol{\theta}\in\Theta\}$是定义在样本空间$\mathcal{X}$上的分布族，其中$\boldsymbol{\theta}=(\theta_1,\ldots,\theta_k)$，$\Theta$为参数空间。若其概率函数$f(x;\boldsymbol{\theta})$可表示成如下形式：

$$f(x;\boldsymbol{\theta})=C(\boldsymbol{\theta})\exp\{\sum_{i=1}^{k}Q_i(\boldsymbol{\theta})T_i(x)\}h(x),$$

则称此分布族为指数型分布族（简称指数族，exponential family），其中k为正整数，$C(\boldsymbol{\theta})$和$Q_i(\boldsymbol{\theta})$ $(i=1,\ldots,k)$都是定义在参数空间Θ上的函数，$h(x)>0$和$T_i(x)$ $(i=1,2,\ldots,k)$都是定义在样本空间$\mathcal{X}$上的函数。

指数族的一个重要性质是族中的所有分布具有共同的支撑集($G(x)=\{x: f(x;\boldsymbol{\theta})>0\}$)称为概率函数$f(x;\boldsymbol{\theta})$的支撑集)。

指数族的一个重要性质是族中的所有分布具有共同的支撑集($G(x) = \{x : f(x; \boldsymbol{\theta}) > 0\}$)称为概率函数$f(x; \boldsymbol{\theta})$的支撑集）。

由定义可见指数族的支持集为$\{x : f(x; \boldsymbol{\theta}) > 0\} = \{x : h(x) > 0\}$与$\boldsymbol{\theta}$无关。

指数族的一个重要性质是族中的所有分布具有共同的支撑集($G(x)=\{x: f(x;\boldsymbol{\theta})>0\}$)称为概率函数$f(x;\boldsymbol{\theta})$的支撑集)。

由定义可见指数族的支持集为$\{x: f(x;\boldsymbol{\theta})>0\}=\{x: h(x)>0\}$与$\boldsymbol{\theta}$无关。

任一分布族若其支撑集与$\boldsymbol{\theta}$有关，则族中分布不再具有共同支撑集，因而不必是指数族。

例 (1.5.1)

设$\boldsymbol{X} = (X_1, \ldots, X_n)$为从正态分布$N(\mu, \sigma^2)$中抽取的简单样本，则样本分布族是指数族。

解

样本$\boldsymbol{X}$的联合密度为

例 (1.5.1)

设$\boldsymbol{X} = (X_1, \ldots, X_n)$为从正态分布$N(\mu, \sigma^2)$中抽取的简单样本，则样本分布族是指数族。

解

样本$\boldsymbol{X}$的联合密度为

$$f(\boldsymbol{x}; \mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

1.5.1 定义与例子

记$\boldsymbol{\theta} = (\mu, \sigma^2)$，则参数空间为$\Theta = \{\boldsymbol{\theta} = (\mu, \sigma^2) : -\infty < \mu < \infty,\ \sigma > 0\}$。

上式可改写为

记$\boldsymbol{\theta} = (\mu, \sigma^2)$，则参数空间为$\Theta = \{\boldsymbol{\theta} = (\mu, \sigma^2) : -\infty < \mu < \infty,\ \sigma > 0\}$。

上式可改写为

$$
\begin{aligned}
f(\boldsymbol{x};\boldsymbol{\theta}) =& (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{n\mu^2}{2\sigma^2}} \exp\left\{\frac{\mu}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n} x_i^2\right\} \\
=&
\end{aligned}
$$

记$\boldsymbol{\theta} = (\mu, \sigma^2)$，则参数空间为$\Theta = \{\boldsymbol{\theta} = (\mu, \sigma^2) : -\infty < \mu < \infty,\ \sigma > 0\}$。上式可改写为

$$
\begin{aligned}
f(\boldsymbol{x}; \boldsymbol{\theta}) =&(2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{n\mu^2}{2\sigma^2}} \exp\left\{ \frac{\mu}{\sigma^2} \sum_{i=1}^{n} x_i - \frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2 \right\} \\
=& C(\boldsymbol{\theta}) \exp\{Q_1(\boldsymbol{\theta}) T_1(\boldsymbol{x}) + Q_2(\boldsymbol{\theta}) T_2(\boldsymbol{x})\} h(\boldsymbol{x}),
\end{aligned}
$$

记$\boldsymbol{\theta} = (\mu, \sigma^2)$，则参数空间为$\Theta = \{\boldsymbol{\theta} = (\mu, \sigma^2) : -\infty < \mu < \infty,\ \sigma > 0\}$。上式可改写为

$$
\begin{aligned}
f(\boldsymbol{x}; \boldsymbol{\theta}) =& (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{n\mu^2}{2\sigma^2}} \exp\left\{ \frac{\mu}{\sigma^2} \sum_{i=1}^{n} x_i - \frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2 \right\} \\
=& C(\boldsymbol{\theta}) \exp\{Q_1(\boldsymbol{\theta}) T_1(\boldsymbol{x}) + Q_2(\boldsymbol{\theta}) T_2(\boldsymbol{x})\} h(\boldsymbol{x}),
\end{aligned}
$$

其中$C(\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{n\mu^2}{2\sigma^2}\}$，$Q_1(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2}$，$Q_2(\boldsymbol{\theta}) = -\frac{1}{2\sigma^2}$，$T_1(\boldsymbol{x}) = \sum\limits_{i=1}^{n} x_i$，$T_2(\boldsymbol{x}) = \sum\limits_{i=1}^{n} x_i^2$，$h(\boldsymbol{x}) = 1$。

1.5.1 定义与例子

记$\boldsymbol{\theta} = (\mu, \sigma^2)$，则参数空间为$\Theta = \{\boldsymbol{\theta} = (\mu, \sigma^2) : -\infty < \mu < \infty,\ \sigma > 0\}$。上式可改写为

$$\begin{aligned} f(\boldsymbol{x}; \boldsymbol{\theta}) =& (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{n\mu^2}{2\sigma^2}} \exp\left\{ \frac{\mu}{\sigma^2} \sum_{i=1}^{n} x_i - \frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2 \right\} \\ =& C(\boldsymbol{\theta}) \exp\{Q_1(\boldsymbol{\theta}) T_1(\boldsymbol{x}) + Q_2(\boldsymbol{\theta}) T_2(\boldsymbol{x})\} h(\boldsymbol{x}), \end{aligned}$$

其中$C(\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{n\mu^2}{2\sigma^2}\}$，$Q_1(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2}$，$Q_2(\boldsymbol{\theta}) = -\frac{1}{2\sigma^2}$，$T_1(\boldsymbol{x}) = \sum\limits_{i=1}^{n} x_i$，$T_2(\boldsymbol{x}) = \sum\limits_{i=1}^{n} x_i^2$，$h(\boldsymbol{x}) = 1$。

由定义1.5.1知上述样本分布族是指数族。

例 (1.5.2)

设$\boldsymbol{X} = (X_1, \ldots, X_n)$为从Gamma分布$\Gamma(\gamma, \lambda)$中抽取的简单样本，则样本分布族是指数族。

例 (1.5.2)

设$\boldsymbol{X} = (X_1, \ldots, X_n)$为从Gamma分布$\Gamma(\gamma, \lambda)$中抽取的简单样本，则样本分布族是指数族。

解

样本$\boldsymbol{X}$的联合密度函数为

例 (1.5.2)

设$\boldsymbol{X}=(X_1,\ldots,X_n)$为从Gamma分布$\Gamma(\gamma,\lambda)$中抽取的简单样本，则样本分布族是指数族。

解

样本$\boldsymbol{X}$的联合密度函数为

$$f(\boldsymbol{x};\gamma,\lambda)=\prod_{i=1}^{n}\left[\frac{\lambda^{\gamma}}{\Gamma(\gamma)}x_i^{\gamma-1}\exp\{-\lambda x_i\}\right],\ x_i>0,\ i=1,\ldots,n.$$

例 (1.5.2)

设$\boldsymbol{X} = (X_1, \ldots, X_n)$为从Gamma分布$\Gamma(\gamma, \lambda)$中抽取的简单样本，则样本分布族是指数族。

解

样本$\boldsymbol{X}$的联合密度函数为

$$f(\boldsymbol{x}; \gamma, \lambda) = \prod_{i=1}^{n} \left[\frac{\lambda^{\gamma}}{\Gamma(\gamma)} x_i^{\gamma-1} \exp\{-\lambda x_i\} \right], \ x_i > 0, \ i = 1, \ldots, n.$$

记$\boldsymbol{\theta} = (\gamma, \lambda)$，则参数空间为$\Theta = \{\boldsymbol{\theta} = (\gamma, \lambda) : \gamma > 0, \lambda > 0\}$。

将上式改写为

将上式改写为

$$f(\boldsymbol{x};\boldsymbol{\theta})=\frac{\lambda^{n\gamma}}{[\Gamma(\gamma)]^n}\exp\left\{-\lambda\sum_{i=1}^{n}x_i+(\gamma-1)\sum_{i=1}^{n}\log x_i\right\}\prod_{i=1}^{n}I_{(0,\infty)}(x_i)$$

1.5.1 定义与例子

将上式改写为

$$f(\boldsymbol{x};\boldsymbol{\theta})=\frac{\lambda^{n\gamma}}{[\Gamma(\gamma)]^n}\exp\left\{-\lambda\sum_{i=1}^n x_i+(\gamma-1)\sum_{i=1}^n\log x_i\right\}\prod_{i=1}^n I_{(0,\infty)}(x_i)$$

其中$C(\boldsymbol{\theta})=\frac{\lambda^{n\gamma}}{[\Gamma(\gamma)]^n}$，$Q_1(\boldsymbol{\theta})=-\lambda$，$Q_2(\boldsymbol{\theta})=\gamma-1$，$T_1(\boldsymbol{x})=\sum_{i=1}^n x_i$，$T_2(\boldsymbol{x})=\sum_{i=1}^n\log x_i$，$h(\boldsymbol{x})=\prod_{i=1}^n I_{(0,\infty)}(x_i)$。

将上式改写为

$$f(\boldsymbol{x};\boldsymbol{\theta}) = \frac{\lambda^{n\gamma}}{[\Gamma(\gamma)]^n}\exp\left\{-\lambda\sum_{i=1}^n x_i + (\gamma-1)\sum_{i=1}^n \log x_i\right\}\prod_{i=1}^n I_{(0,\infty)}(x_i)$$

其中$C(\boldsymbol{\theta}) = \frac{\lambda^{n\gamma}}{[\Gamma(\gamma)]^n}$，$Q_1(\boldsymbol{\theta}) = -\lambda$，$Q_2(\boldsymbol{\theta}) = \gamma - 1$，$T_1(\boldsymbol{x}) = \sum_{i=1}^n x_i$，$T_2(\boldsymbol{x}) = \sum_{i=1}^n \log x_i$，$h(\boldsymbol{x}) = \prod_{i=1}^n I_{(0,\infty)}(x_i)$。

由定义1.5.1知上述样本分布族是指数族。

例 (1.5.3)

二项分布族$\{B(n, \theta) : 0 < \theta < 1\}$是指数族。

例 (1.5.3)

二项分布族$\{B(n, \theta) : 0 < \theta < 1\}$是指数族。

解

设$X \sim B(n, \theta)$，其概率函数为

例 (1.5.3)

二项分布族$\{B(n,\theta):0<\theta<1\}$是指数族。

解

设$X\sim B(n,\theta)$，其概率函数为

$$\begin{aligned} p(x;\theta)=&P_\theta(X=x)=\binom{n}{x}\theta^x(1-\theta)^{n-x}\\ =&\binom{n}{x}\left(\frac{\theta}{1-\theta}\right)^x(1-\theta)^n,\ x=0,1,\ldots,n, \end{aligned}$$

例 (1.5.3)

二项分布族$\{B(n,\theta):0<\theta<1\}$是指数族。

解

设$X\sim B(n,\theta)$，其概率函数为

$$\begin{aligned}p(x;\theta)=&P_\theta(X=x)=\binom{n}{x}\theta^x(1-\theta)^{n-x}\\=&\binom{n}{x}\left(\frac{\theta}{1-\theta}\right)^x(1-\theta)^n,\ x=0,1,\ldots,n,\end{aligned}$$

其样本空间为$\mathcal{X}=\{0,1,\ldots,n\}$，参数空间为$\Theta=\{\theta:0<\theta<1\}$。

将上式改写为

将上式改写为

$$
\begin{aligned}
p(x;\theta) =&(1-\theta)^n \exp\left\{x\ln\frac{\theta}{1-\theta}\right\}\binom{n}{x} \\
=&C(\theta)\exp\{Q_1(\theta)T_1(x)\}h(x),
\end{aligned}
$$

将上式改写为

$$
\begin{aligned}
p(x;\theta) =&(1-\theta)^n \exp\left\{x \ln \frac{\theta}{1-\theta}\right\}\binom{n}{x} \\
=&C(\theta)\exp\{Q_1(\theta)T_1(x)\}h(x),
\end{aligned}
$$

其中$C(\theta)=(1-\theta)^n$，$Q_1(\theta)=\ln\frac{\theta}{1-\theta}$，$T_1(x)=x$，$h(x)=\binom{n}{x}$。

将上式改写为

$$
\begin{aligned}
p(x;\theta) =& (1-\theta)^n \exp\left\{x \ln \frac{\theta}{1-\theta}\right\} \binom{n}{x} \\
=& C(\theta) \exp\{Q_1(\theta) T_1(x)\} h(x),
\end{aligned}
$$

其中$C(\theta) = (1-\theta)^n$，$Q_1(\theta) = \ln \frac{\theta}{1-\theta}$，$T_1(x) = x$，$h(x) = \binom{n}{x}$。

由定义1.5.1知上述样本分布族是指数族。

例 (1.5.4)

Poisson分布族$\{P(\theta) : \theta > 0\}$是指数族。

例 (1.5.4)

Poisson分布族$\{P(\theta):\theta>0\}$是指数族。

解

设X服从泊松分布$P(\theta)$，其概率函数为

例 (1.5.4)

Poisson分布族$\{P(\theta) : \theta > 0\}$是指数族。

解

设X服从泊松分布$P(\theta)$，其概率函数为

$$
\begin{aligned}
p(x;\theta) =& P_\theta(X = x) = \frac{e^{-\theta}\theta^x}{x!} = \frac{e^{-\theta}\exp\{x\ln\theta\}}{x!} \\
=& C(\theta)\exp\{Q_1(\theta)T_1(x)\}h(x),\ x = 0, 1, \ldots,
\end{aligned}
$$

例 (1.5.4)

Poisson分布族$\{P(\theta):\theta>0\}$是指数族。

解

设X服从泊松分布$P(\theta)$，其概率函数为

$$
\begin{aligned}
p(x;\theta) =& P_\theta(X=x) = \frac{e^{-\theta}\theta^x}{x!} = \frac{e^{-\theta}\exp\{x\ln\theta\}}{x!} \\
=& C(\theta)\exp\{Q_1(\theta)T_1(x)\}h(x),\ x=0,1,\ldots,
\end{aligned}
$$

其中样本空间为$\mathcal{X}=\{0,1,\ldots,\}$，参数空间为$\Theta=\{\theta:\theta>0\}$，
$C(\theta)=e^{-\theta}$，$Q_1(\theta)=\ln\theta$，$T_1(x)=x$，$h(x)=\frac{1}{x!}$。

例 (1.5.4)

Poisson分布族$\{P(\theta):\theta>0\}$是指数族。

解

设X服从泊松分布$P(\theta)$，其概率函数为

$$\begin{aligned} p(x;\theta) =& P_\theta(X=x)=\frac{e^{-\theta}\theta^x}{x!}=\frac{e^{-\theta}\exp\{x\ln\theta\}}{x!} \\ =& C(\theta)\exp\{Q_1(\theta)T_1(x)\}h(x),\ x=0,1,\dots, \end{aligned}$$

其中样本空间为$\mathcal{X}=\{0,1,\dots,\}$，参数空间为$\Theta=\{\theta:\theta>0\}$，$C(\theta)=e^{-\theta}$，$Q_1(\theta)=\ln\theta$，$T_1(x)=x$，$h(x)=\frac{1}{x!}$。

由定义1.5.1知上述样本分布族是指数族。

例 (1.5.5)

均匀分布族$\{U(0,\theta):\theta>0\}$和双参数指数分布不是指数族。

解

例 (1.5.5)

均匀分布族$\{U(0,\theta):\theta>0\}$和双参数指数分布不是指数族。

解

设X服从均匀分布$U(0,\theta)$，其密度函数为$f(x;\theta)=I_{0,\theta}(x)$，支撑集为$\{x:f(x;\theta)>0\}=\{x:0<x<\theta\}$与参数$\theta$有关，不满足指数族的定义，因此均匀分布族不是指数族。

若设X服从双参数指数分布族，其密度函数为

若设X服从双参数指数分布族，其密度函数为

$$
\begin{aligned}
p(x;\mu,\sigma) =& \frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\} I_{[x>\mu]} \\
=&
\end{aligned}
$$

若设X服从双参数指数分布族，其密度函数为

$$
\begin{aligned}
p(x;\mu,\sigma) =& \frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\} I_{[x>\mu]} \\
=& \frac{1}{\sigma}\exp\{\frac{\mu}{\sigma}\}\exp\{-\frac{x}{\sigma}\}I_{[x>\mu]}, \ -\infty<\mu<\infty, \ \sigma>0,
\end{aligned}
$$

若设X服从双参数指数分布族，其密度函数为

$$
\begin{aligned}
p(x;\mu,\sigma) =& \frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\} I_{[x>\mu]} \\
=& \frac{1}{\sigma}\exp\{\frac{\mu}{\sigma}\}\exp\{-\frac{x}{\sigma}\} I_{[x>\mu]},\ -\infty<\mu<\infty,\ \sigma>0,
\end{aligned}
$$

它的支撑集为$\{x: p(x;\mu,\sigma)>0\} = \{x: x>\mu\} = (\mu,\infty)$与未知参数$\mu$有关，因此双参数指数分布不是指数族。

若设X服从双参数指数分布族，其密度函数为

$$
\begin{aligned}
p(x;\mu,\sigma) =& \frac{1}{\sigma}\exp\left\{-\frac{x-\mu}{\sigma}\right\} I_{[x>\mu]} \\
=& \frac{1}{\sigma}\exp\{\frac{\mu}{\sigma}\}\exp\{-\frac{x}{\sigma}\}I_{[x>\mu]},\ -\infty<\mu<\infty,\ \sigma>0,
\end{aligned}
$$

它的支撑集为$\{x : p(x;\mu,\sigma) > 0\} = \{x : x > \mu\} = (\mu,\infty)$与未知参数$\mu$有关，因此双参数指数分布不是指数族。

但若μ已知，如$\mu = 0$，则单参数指数分布族$Exp(1/\sigma)$属于指数族。

将样本分布族表示为指数族的自然形式：在指数族的定义$C(\boldsymbol{\theta})\exp\{\sum_{i=1}^{k}Q_i(\boldsymbol{\theta})T_i(x)\}h(x)$中，若用$\varphi_i$代替$Q_i(\boldsymbol{\theta})$，而将$C(\boldsymbol{\theta})$表示成$\boldsymbol{\varphi}$的函数$C^*(\boldsymbol{\varphi})$，$\boldsymbol{\varphi}=(\varphi_1,\ldots,\varphi_k)$，则指数族的表达式变为$C^*(\boldsymbol{\varphi})\exp\{\sum_{i=1}^{k}\varphi_i T_i(x)\}h(x)$，再改$\varphi_i$为$\theta_i$，$i=1,\ldots,k$，得到如下指数族的自然形式（或称为标准形式）的定义。

1.5.2 指数族的自然形式及自然参数空间

定义 (1.5.2)

如果指数族有下列形式：

$$f(x,\boldsymbol{\theta}) = C^*(\boldsymbol{\theta}) \exp\{\sum_{i=1}^{k} \theta_i T_i(x)\} h(x)$$

则称它为指数族的自然形式（natural form）。此时集合

$$\Theta^* = \left\{ (\theta_1, \ldots, \theta_k) : \int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} \theta_i T_i(x)\} h(x) dx < \infty \right\}$$

称为自然参数空间。

例 (1.5.6)

将例1.5.1中样本分布族$N(\mu,\sigma^2)$表示为指数族的自然形式，并求其自然参数空间。

解 由例1.5.1可知该样本分布族的指数族形式为

例 (1.5.6)

将例1.5.1中样本分布族$N(\mu,\sigma^2)$表示为指数族的自然形式，并求其自然参数空间。

解 由例1.5.1可知该样本分布族的指数族形式为

$$
\begin{aligned}
f(\boldsymbol{x};\boldsymbol{\theta}) =&(2\pi\sigma^2)^{-\frac{n}{2}}\exp\{-\frac{n\mu^2}{2\sigma^2}\}\exp\{\frac{\mu}{\sigma^2}\sum_{i=1}^n x_i-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2\}\\
=&C(\boldsymbol{\theta})\exp\{Q_1(\boldsymbol{\theta})T_1(\boldsymbol{x})+Q_2(\boldsymbol{\theta})T_2(\boldsymbol{x})\}h(\boldsymbol{x}).
\end{aligned}
$$

例 (1.5.6)

将例1.5.1中样本分布族$N(\mu,\sigma^2)$表示为指数族的自然形式，并求其自然参数空间。

解 由例1.5.1可知该样本分布族的指数族形式为

$$\begin{aligned} f(\boldsymbol{x};\boldsymbol{\theta}) =&(2\pi\sigma^2)^{-\frac{n}{2}}\exp\{-\frac{n\mu^2}{2\sigma^2}\}\exp\{\frac{\mu}{\sigma^2}\sum_{i=1}^n x_i-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2\} \\ =&C(\boldsymbol{\theta})\exp\{Q_1(\boldsymbol{\theta})T_1(\boldsymbol{x})+Q_2(\boldsymbol{\theta})T_2(\boldsymbol{x})\}h(\boldsymbol{x}). \end{aligned}$$

令$\varphi_1=Q_1(\boldsymbol{\theta})$，$\varphi_2=Q_2(\boldsymbol{\theta})$，解出$\mu=-\frac{\varphi_1}{2\varphi_2}$，$\sigma^2=-\frac{1}{2\varphi_2}$，从而$\frac{\mu^2}{\sigma^2}=-\frac{\varphi_1^2}{2\varphi_2}$。

因此$C(\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{n\mu^2}{2\sigma^2}\}$变为$C^*(\boldsymbol{\varphi}) = (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\}$，$\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$。

因此$C(\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{n\mu^2}{2\sigma^2}\}$变为$C^*(\boldsymbol{\varphi}) = (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\}$，$\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$。 可获得自然形式为

因此$C(\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{n\mu^2}{2\sigma^2}\}$变为$C^*(\boldsymbol{\varphi}) = (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\}$，$\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$。可获得自然形式为

$$
\begin{aligned}
f(\boldsymbol{x}; \boldsymbol{\varphi}) =& (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\} \exp\{\varphi_1 T_1(\boldsymbol{x}) + \varphi_2 T_2(\boldsymbol{x})\} \\
=& C^*(\boldsymbol{\varphi}) \exp\{\varphi_1 T_1(\boldsymbol{x}) + \varphi_2 T_2(\boldsymbol{x})\} h(\boldsymbol{x}).
\end{aligned}
$$

因此$C(\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{n\mu^2}{2\sigma^2}\}$变为$C^*(\boldsymbol{\varphi}) = (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\}$，$\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$。 可获得自然形式为

$$\begin{aligned} f(\boldsymbol{x}; \boldsymbol{\varphi}) =& (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\} \exp\{\varphi_1 T_1(\boldsymbol{x}) + \varphi_2 T_2(\boldsymbol{x})\} \\ =& C^*(\boldsymbol{\varphi}) \exp\{\varphi_1 T_1(\boldsymbol{x}) + \varphi_2 T_2(\boldsymbol{x})\} h(\boldsymbol{x}). \end{aligned}$$

其自然参数空间为

因此$C(\boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{n\mu^2}{2\sigma^2}\}$变为$C^*(\boldsymbol{\varphi}) = (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\}$，$\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$。 可获得自然形式为

$$
\begin{aligned}
f(\boldsymbol{x}; \boldsymbol{\varphi}) =& (-\frac{\pi}{\varphi_2})^{-\frac{n}{2}} \exp\{\frac{n\varphi_1^2}{4\varphi_2}\} \exp\{\varphi_1 T_1(\boldsymbol{x}) + \varphi_2 T_2(\boldsymbol{x})\} \\
=& C^*(\boldsymbol{\varphi}) \exp\{\varphi_1 T_1(\boldsymbol{x}) + \varphi_2 T_2(\boldsymbol{x})\} h(\boldsymbol{x}).
\end{aligned}
$$

其自然参数空间为

$$
\Theta^* = \{\boldsymbol{\varphi} = (\varphi_1, \varphi_2) : -\infty < \varphi_1 < \infty,\ -\infty < \varphi_2 < 0\}.
$$

例 (1.5.7)

将例1.5.3中二项分布族$\{B(n,\theta):0<\theta<1\}$表示为指数族的自然形式，并求其自然参数空间。

解 由例1.5.3可知该样本分布族的指数族形式为

例 (1.5.7)

将例1.5.3中二项分布族$\{B(n,\theta):0<\theta<1\}$表示为指数族的自然形式，并求其自然参数空间。

解 由例1.5.3可知该样本分布族的指数族形式为

$$
\begin{aligned}
f(x;\theta) =&(1-\theta)^n \exp\{\ln \frac{\theta}{1-\theta} x\}\binom{n}{x},\\
=&C(\theta)\exp\{Q(\theta)T(x)\}h(x),\ x=0,1,\ldots,n.
\end{aligned}
$$

令$\varphi=\ln\frac{\theta}{1-\theta}$，解出$\theta=\frac{1}{1+e^{-\varphi}}$。因此$C(\theta)=(1-\theta)^n$变为$C^*(\varphi)=(1+e^{\varphi})^{-n}$。

故可获得自然形式为

$$
\begin{aligned}
f(x;\varphi) =&(1+e^{\varphi})^{-n}\exp\{\varphi x\}\binom{n}{x} \\
=&C(\varphi)\exp\{\varphi x\}h(x).
\end{aligned}
$$

其自然参数空间为

$$\Theta^* = \{\varphi : -\infty < \varphi < \infty\} = (-\infty, \infty).$$

定理 (1.5.1)

在指数族的自然形式下，自然参数空间为凸集。

定理 (1.5.1)

在指数族的自然形式下，自然参数空间为凸集。

证明

设$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$为$k$维向量。任给$\boldsymbol{\theta}^{(1)} = (\theta_1^1, \ldots, \theta_k^1)$，$\boldsymbol{\theta}^{(0)} = (\theta_1^0, \ldots, \theta_k^0)$皆属于自然参数空间$\Theta^*$。

定理 (1.5.1)

在指数族的自然形式下，自然参数空间为凸集。

证明

设$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$为$k$维向量。任给$\boldsymbol{\theta}^{(1)} = (\theta_1^1, \ldots, \theta_k^1)$，$\boldsymbol{\theta}^{(0)} = (\theta_1^0, \ldots, \theta_k^0)$皆属于自然参数空间$\Theta^*$。

设$0 < \alpha < 1$，令$\boldsymbol{\theta} = \alpha\boldsymbol{\theta}^{(1)} + (1-\alpha)\boldsymbol{\theta}^{(0)}$，

即$\theta_i = \alpha\theta_i^1 + (1-\alpha)\theta_i^0$，$i = 1, \ldots, k$，则有

$$\int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} \theta_i T_i(x)\} h(x) dx$$

$$
\begin{aligned}
&\int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} \theta_i T_i(x)\} h(x) dx \\
= &\int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} [\alpha \theta_i^{(1)} + (1-\alpha)\theta_i^{(0)}] T_i(x)\} h(x) dx
\end{aligned}
$$

$$
\begin{aligned}
&\int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} \theta_i T_i(x)\} h(x) dx \\
=&\int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} [\alpha\theta_i^{(1)} + (1-\alpha)\theta_i^{(0)}] T_i(x)\} h(x) dx \\
=&\int_{\mathcal{X}} \left[\exp\{\sum_{i=1}^{k} \theta_i^{(1)} T_i(x)\}\right]^{\alpha} \left[\exp\{\sum_{i=1}^{k} \theta_i^{(0)} T_i(x)\}\right]^{1-\alpha} h(x) dx
\end{aligned}
$$

由Holder不等式有$EX^\alpha Y^{1-\alpha} \leq (EX)^\alpha (EY)^{1-\alpha}$，这里$0 < \alpha < 1$，则上式

由Holder不等式有$EX^\alpha Y^{1-\alpha} \leq (EX)^\alpha (EY)^{1-\alpha}$，这里$0 < \alpha < 1$，则上式

$$\leq \left[\int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} \theta_i^{(1)} T_i(x)\} h(x) dx\right]^{\alpha} \times \left[\int_{\mathcal{X}} \exp\{\sum_{i=1}^{k} \theta_i^{(0)} T_i(x)\} h(x) dx\right]^{1-\alpha},$$

由Holder不等式有$EX^\alpha Y^{1-\alpha} \le (EX)^\alpha (EY)^{1-\alpha}$，这里$0<\alpha<1$，则上式

$$\le \left[\int_{\mathcal{X}} \exp\{\sum_{i=1}^k \theta_i^{(1)} T_i(x)\} h(x) dx\right]^\alpha \times \left[\int_{\mathcal{X}} \exp\{\sum_{i=1}^k \theta_i^{(0)} T_i(x)\} h(x) dx\right]^{1-\alpha},$$

所以不等式右边两个积分存在，从而得到$\boldsymbol{\theta} \in \Theta^*$，这说明了自然参数空间$\Theta^*$是凸集。

1.5.3 指数族的性质

指数族具有良好的分析性质。

指数族具有良好的分析性质。

定理 (1.5.2)

设指数族的自然形式中，自然参数空间有内点，其内点集为Θ_0。

设$g(x)$为任一实函数，使得积分

指数族具有良好的分析性质。

定理 (1.5.2)

设指数族的自然形式中，自然参数空间有内点，其内点集为Θ_0。

设$g(x)$为任一实函数，使得积分

$$G(\boldsymbol{\theta}) = \int_{\mathcal{X}} g(x) \exp\{\sum_{j=1}^{k} \theta_j T_j(x)\} h(x) dx$$

在Θ_0内存在且有限，则$G(\boldsymbol{\theta})$的任意阶偏导数在Θ_0内存在且可在积分号下求得，即

1.5.3 指数族的性质

定理 (1.5.2)

$$\frac{\partial^m G(\boldsymbol{\theta})}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}} = \int_{\mathcal{X}} \frac{\partial^m}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}}\left[g(x)\exp\{\sum_{j=1}^{k}\theta_j T_j(x)\}h(x)\right]dx,$$

定理 (1.5.2)

$$\frac{\partial^m G(\boldsymbol{\theta})}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}} = \int_{\mathcal{X}} \frac{\partial^m}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}} \left[g(x)\exp\{\sum_{j=1}^{k}\theta_j T_j(x)\}h(x) \right] dx,$$

其中$\sum\limits_{j=1}^{k} m_j = m$，即对$G(\boldsymbol{\theta})$关于$\boldsymbol{\theta}$的任意阶偏导数可在积分下求得。

定理 (1.5.2)

$$\frac{\partial^m G(\boldsymbol{\theta})}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}} = \int_{\mathcal{X}} \frac{\partial^m}{\partial\theta_1^{m_1}\cdots\partial\theta_k^{m_k}}\left[g(x)\exp\{\sum_{j=1}^{k}\theta_j T_j(x)\}h(x)\right]dx,$$

其中$\sum\limits_{j=1}^{k} m_j = m$，即对$G(\boldsymbol{\theta})$关于$\boldsymbol{\theta}$的任意阶偏导数可在积分下求得。

证明：略。