样本及抽样分布

李孟棠，朱彬

中山大学智能工程学院

limt29@mail.sysu.edu.cn, zhub26@mail.sysu.edu.cn

2021 年 x 月 x 日

主要内容

下面的课程进入数理统计 (mathematical statistics) 部分。

统计学以概率论为理论基础，根据实验或观察得到的数据来研究随机现象，对研究对象的规律、性质、特点做出推断，称为统计推断 (statistical inference)。

例如：概率论研究的随机变量，其分布均假设已知。在此前提下我们研究它的性质，例如数字特征（数学期望、方差等）和随机变量函数的分布等。在数理统计中，随机变量的分布未知或部分未知，人们通过对随机变量进行观察、对所得数据进行分析，从而对随机变量的分布做出推断。

主要内容

几个名词：

- 一个试验全部可能的观察值称为**总体**，
- 每一个可能的观察值称为**个体**，
- 总体中所包含的个体数量称为总体的**容量**，
- 容量有限的称为**有限总体**，反之称为**无限总体**。

一个总体对应于一个随机变量 X，笼统地称为总体 X；每一个个体是随机变量的值。

抽样：

- 在相同条件下，对随机变量 X 进行 n 次**独立**、**重复**的观察，并将结果记为 $X_1, X_2, \ldots, X_n$。
- 我们把 $X_1, X_2, \ldots, X_n$ 视为 n 个相互独立、与 X 具有相同分布的 (i.i.d.) **随机变量**，称为总体 X 的一个简单随机样本，简称**样本**，n 为样本容量。
- 当 n 次观察完成时，我们得到一组实数 $x_1, x_2, \ldots, x_n$，称为随机变量 $X_1, X_2, \ldots, X_n$ 的**样本值**或 n 个**独立观察值**。

我们也可以把样本看作一个随机向量 $(X_1, X_2, \ldots, X_n)$，一次观察的样本值写作 $(x_1, x_2, \ldots, x_n)$。若做另一次观察得到样本值 $(y_1, y_2, \ldots, y_n)$，一般来说两次观察值不同。

由样本的定义得：如果 $X_1, X_2, \ldots, X_n$ 是随机变量 X 的 i.i.d. 样本，每个 X_i 的分布函数均为 F，那么 $(X_1, X_2, \ldots, X_n)$ 的分布函数为

$$F^*(x_1, \ldots, x_n) = \prod_{i=1}^{n} F(x_i).$$

如果 X 的具有概率密度 f，那么 $(X_1, X_2, \ldots, X_n)$ 的概率密度为

$$f^*(x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i).$$

关于直方图和箱线图，请自行阅读课本第六章第二节。

要求：知晓 p 分位数和疑似异常值的定义。

参考 Matlab 的`boxplot` 命令。

主要内容

构造样本的适当函数进行统计推断。

Definition

设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的一个样本，$g(X_1, \dots, X_n)$ 是样本的函数。若 g 中不含未知参数，则称 $g(X_1, \dots, X_n)$ 为一个统计量 (statistic)。

显然统计量 $g(X_1, \dots, X_n)$ 是一个随机变量。设 $x_1, \dots, x_n$ 是 $X_1, \dots, X_n$ 的样本值，则称 $g(x_1, \dots, x_n)$ 是 $g(X_1, \dots, X_n)$ 的观察值。

几个常用的统计量：

样本均值

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

样本方差

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

样本标准差

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

样本 k 阶（原点）矩

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

样本 k 阶中心矩

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

把样本 $X_1,\ldots,X_n$ 的观察值 $x_1,\ldots,x_n$ 带入上面各式，即得到这些统计量的观察值 $\bar{x},s^2,s,a_k,b_k$。

如果总体 X 的 k 阶矩存在，记作 $\mu_k := \mathbb{E}(X^k)$，那么当 $n \to \infty$ 时，我们有 $A_k \xrightarrow{P} \mu_k$， $k = 1, 2, \ldots$ 。

原因： $X_1, \ldots, X_n$ 是 X 的 i.i.d. 样本，所以 $X_1^k, \ldots, X_n^k$ 是 X^k 的 i.i.d. 样本。因此，

$$E(X_1^k) = \cdots = E(X_n^k) = \mu_k.$$

由 Khinchin 大数定律知，

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{P} \mu_k, \quad k = 1, 2, \ldots.$$

进而由依概率收敛的性质可知

$$g(A_1, \ldots, A_k) \xrightarrow{P} g(\mu_1, \ldots, \mu_k),$$

其中 g 为连续函数。

对应总体 X 分布函数 $F(x)$ 的统计量——经验分布函数

Definition

设 $x_1, \ldots, x_n$ 是来自分布函数为 $F(x)$ 的总体 X 的样本观察值。X 的经验分布函数 (empirical distribution function)，记作 $F_n(x)$，定义为样本观察值中小于或等于指定值 x 所占的比率，即

$$F_n(x) = \frac{\#\{x_i \le x\}}{n}, \quad x \in \mathbb{R}.$$

按定义，给定样本观察值 $x_1, \ldots, x_n$，$F_n(x)$ 满足分布函数的三个条件：

1. $F_n(x)$ 是 x 的非减函数，
2. $0 \le F_n(x) \le 1$，且 $F_n(-\infty) = 0$，$F_n(\infty) = 1$，
3. F_n 右连续。

当 $x_1, \ldots, x_n$ 各不相同时，$F_n(x)$ 是一离散型随机变量的分布函数，该随机变量以等概率 $1/n$ 取值 $x_1, \ldots, x_n$。

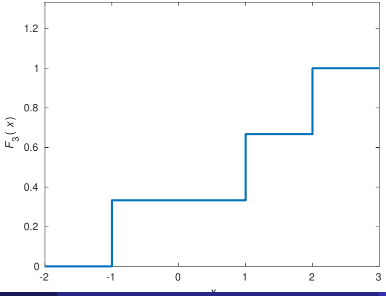
一般而言，给定总体 X 的 n 个样本观察值 $x_1, \ldots, x_n$，先对它们排序并重新编号为

$$x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)},$$

则经验分布函数可写为

$$F_n(x) = \begin{cases} 0, & x < x_{(1)} \\ k/n, & x_{(k)} \le x < x_{(k+1)}, \quad k = 1, 2, \ldots, n-1, \\ 1, & x \ge x_{(n)}. \end{cases}$$

例：设总体 X 有样本观察值 $x_{(1)} = -1$，$x_{(2)} = 1$，$x_{(3)} = 2$，则经验分布函数如图所示。



给定样本观察值 $x_1, \ldots, x_n$，经验分布函数 $F_n(x; x_1, \ldots, x_n)$ 是以样本观察值为参数、以 $x \in \mathbb{R}$ 为变量的函数。

以随机变量 $X_1, \ldots, X_n$ 代替它们的样本值，可知对于任意给定的实数 x，$F_n(x; X_1, \ldots, X_n)$ 是一个随机变量。

事实上，$g(x, e) := F_n(x; X_1(e), \ldots, X_n(e))$ 是定义在 $R \times S$ 上的二元函数，可以看作一个广义的统计量。

定理 (Glivenko-Cantelli, 1933)

设 $X_1, \ldots, X_n$ 是来自以 $F(x)$ 为分布函数的总体 X 的样本，$F_n(x)$ 是经验分布函数，则有

$$P\left\{\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = 0\right\} = 1.$$

证明略。

定理解读：当样本容量 n 充分大时，用总体 X 的 i.i.d. 样本 $X_1, \ldots, X_n$ 构造的经验分布函数 $F_n(x)$ 能够很好地逼近总体分布函数 $F(x)$。具体而言，我们有

$$\|F_n - F\|_\infty := \sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \xrightarrow{\text{a.s.}} 0,$$

即几乎肯定收敛。

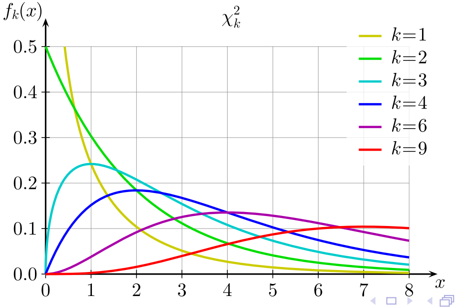
χ^2 分布

设 $X_1, X_2, \ldots, X_n$ 是来自标准正态总体 $N(0,1)$ 的样本，则称统计量

$$Q = X_1^2 + X_2^2 + \cdots + X_n^2$$

服从自由度为 n（独立变量个数）的 χ^2 分布，记为 $Q \sim \chi^2(n)$。其概率密度函数为

$$f(y;n) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2}, & y > 0, \\ 0, & \text{其他}. \end{cases}$$



χ^2 概率密度的推导：利用和 gamma 分布的联系

注意到 $\chi^2(1)$ 分布即为 $\Gamma(\frac{1}{2},2)$ 分布。现有 $X_i \sim N(0,1)$，故由定义知 $X_i^2 \sim \chi^2(1)$，即 $X_i^2 \sim \Gamma(\frac{1}{2},2)$，$i=1,\ldots,n$。再由 $X_1,\ldots,X_n$ 为 i.i.d. 知 $X_1^2,\ldots,X_n^2$ 为 i.i.d.，从而由 gamma 分布的可加性得到

$$Q=\sum_{i=1}^{n} X_i^2 \sim \Gamma(\tfrac{n}{2},2),$$

即为 $\chi^2(n)$ 分布的概率密度。

χ^2 分布的性质

- 可加性。设 $Q_1 \sim \chi^2(n_1)$，$Q_2 \sim \chi^2(n_2)$，且 Q_1, Q_2 相互独立，则有

$$Q_1+Q_2 \sim \chi^2(n_1+n_2).$$

  证明：利用 gamma 分布的可加性，留作练习题。

χ^2 分布的性质（续）

- 数学期望和方差。若 $Q \sim \chi^2(n)$，则有

$$\mathbb{E}(Q) = n, \quad D(Q) = 2n.$$

证明：因为对于 $i = 1, \ldots, n$ 有 $X_i \sim N(0,1)$，所以

$$\begin{aligned} \mathbb{E}(X_i^2) = D(X_i) = 1, \\ D(X_i^2) = \mathbb{E}(X_i^4) - [\mathbb{E}(X_i^2)]^2 = 3 - 1 = 2, \end{aligned}$$

其中第二个式子用到了正态分布的四阶矩。因此，

$$\begin{aligned} \mathbb{E}(Q) = \mathbb{E}\left(\sum_{i=1}^n X_i^2\right) = \sum_{i=1}^n \mathbb{E}\left(X_i^2\right) = n, \\ D(Q) = D\left(\sum_{i=1}^n X_i^2\right) = \sum_{i=1}^n D\left(X_i^2\right) = 2n. \end{aligned}$$

χ^2 分布的性质（续）

- 上分位数。对于给定的正数 $\alpha \in (0,1)$，满足条件

$$P\{Q > \chi_\alpha^2(n)\} = \int_{\chi_\alpha^2(n)}^{\infty} f(y;n)\mathrm{d}y = \alpha$$

的实数 $\chi_\alpha^2(n)$ 称为 $\chi^2(n)$ 分布的上 α 分位数，可查表得到具体数值。Fisher 证明：当 n 充分大时有如下近似公式

$$\chi_\alpha^2(n) \approx \frac{1}{2}(z_\alpha + \sqrt{2n+1})^2,$$

其中 z_α 为标准正态分布的上 α 分位数。

t 分布

英文全称为 Student's t-distribution。设 $X \sim N(0,1)$，$Y \sim \chi^2(n)$，且 X, Y 相互独立，则称随机变量

$$T = \frac{X}{\sqrt{Y/n}}$$

服从自由度为 n 的 t 分布，记作 $T \sim t(n)$。其概率密度函数为

$$h(t;n) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad t \in \mathbb{R}.$$

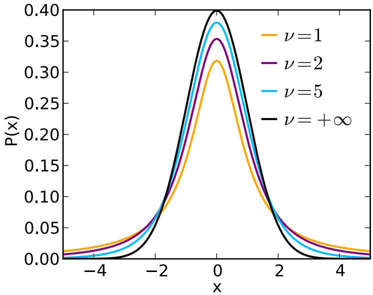
证明略。显然 $h(t;n)$ 是关于 t 的偶函数。

命名“学生”来自英国统计学家、化学家、啤酒酿造师 William Sealy Gosset 以笔名 Student 在 1908 年发表于 Biometrika 的论文，进一步由 Ronald Fisher 进行普及并采用字母 t 表示检验值。

命题：当自由度 $n \to \infty$ 时，t 分布趋向于标准正态分布。

证明：固定 $t \in \mathbb{R}$，考察极限

$$\begin{aligned}
&\lim_{n\to\infty} h(t;n) \\
&= \lim_{n\to\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \\
&= \lim_{n\to\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \cdot \lim_{n\to\infty} \left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}
\end{aligned}$$



不难验证第二项极限等于 $e^{-t^2/2}$（练习题），因此只需证明第一项极限等于 $1/\sqrt{2\pi}$。此时需要利用描述 gamma 函数渐近性质的 Stirling 公式

$$\Gamma(z) \sim \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z, \quad z \in \mathbb{R},\ z \to \infty.$$

计算过程参考https://math.stackexchange.com/questions/3240536/convergence-of-students-t-distribution-to-a-standard-normal.

t 分布的上分位数：对于给定的 $\alpha \in (0,1)$，满足条件

$$P\{T > t_\alpha(n)\} = \int_{t_\alpha(n)}^{\infty} h(t;n)\mathrm{d}t = \alpha$$

的实数 $t_\alpha(n)$ 称为 $t(n)$ 分布的上 α 分位数。

由上述定义和 $h(t;n)$ 函数图象的对称性知

$$t_{1-\alpha}(n) = -t_\alpha(n).$$

当 $n > 45$ 时，对于常用的 α 值，我们采用近似

$$t_\alpha(n) \approx z_\alpha,$$

其中 z_α 为标准正态分布的上 α 分位数。

F 分布

设 $U \sim \chi^2(n_1)$，$V \sim \chi^2(n_2)$，且 U, V 相互独立，则称随机变量

$$X = \frac{U/n_1}{V/n_2}$$

服从自由度为 (n_1, n_2) 的 F 分布，记为 $X \sim F(n_1, n_2)$。其概率密度函数为

$$\psi(y; n_1, n_2) = \begin{cases} \dfrac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} y^{\frac{n_1}{2}-1}}{\mathrm{B}(\frac{n_1}{2}, \frac{n_2}{2})(1+\frac{n_1}{n_2}y)^{\frac{n_1+n_2}{2}}}, & y > 0, \\ 0, & \text{其他}, \end{cases}$$

其中，B 表示 beta 函数，对于实部为正 (即 $\mathrm{Re}(x) > 0$，$\mathrm{Re}(y) > 0$) 的复数 x, y，其定义为

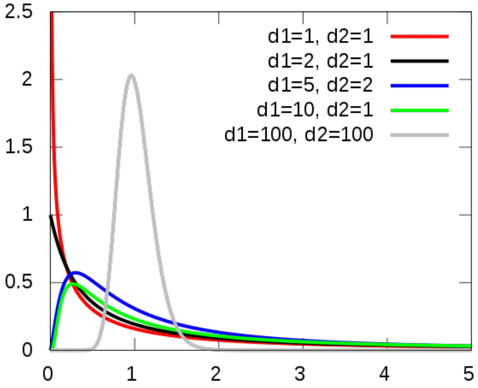
$$\mathrm{B}(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t.$$

利用 beta 函数和 gamma 函数的如下联系

$$B(x,y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)},$$

我们即可恢复书上 F 分布的概率密度表达式。

概率密度函数的图象如下（其中 (d_1, d_2) 表示自由度）:



由定义可知，若 $X \sim F(n_1, n_2)$，则有

$$\frac{1}{X} \sim F(n_2, n_1).$$

F 分布的上分位数：对于给定的 $\alpha \in (0, 1)$，满足条件

$$P\{X > F_\alpha(n_1, n_2)\} = \int_{F_\alpha(n_1,n_2)}^{\infty} \psi(y; n_1, n_2)\mathrm{d}y = \alpha$$

的实数 $F_\alpha(n_1, n_2)$ 称为 $F(n_1, n_2)$ 分布的上 α 分位数。它满足如下重要性质：

$$F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)}. \tag{1}$$

该性质常被用来求 F 分布表中未列出的上 α 分位数。

性质 (1) 的证明.

设 $X \sim F(n_1, n_2)$，则根据定义有

$$1-\alpha = P\{X > F_{1-\alpha}(n_1,n_2)\} = P\left\{\frac{1}{X} < \frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}$$
$$= 1 - P\left\{\frac{1}{X} > \frac{1}{F_{1-\alpha}(n_1,n_2)}\right\}.$$

注意花括号中等号事件的概率为零。因此

$$P\left\{\frac{1}{X} > \frac{1}{F_{1-\alpha}(n_1,n_2)}\right\} = \alpha.$$

又因为 $\frac{1}{X} \sim F(n_2, n_1)$，所以

$$P\left\{\frac{1}{X} > F_\alpha(n_2,n_1)\right\} = \alpha.$$

比较两式即得到性质 (1)。 $\square$

设总体 X 的分布未知，但具有有限的均值 μ 和方差 σ^2。再设 $X_1, \ldots, X_n$ 是 X 的样本，$\bar{X}, S^2$ 分别为样本均值和样本方差，则有

$$\mathbb{E}(\bar{X}) = \mu, \quad D(\bar{X}) = \frac{\sigma^2}{n}.$$

$$\begin{aligned}
\mathbb{E}(S^2) = \mathbb{E}\left[\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)\right] &= \frac{1}{n-1}\left[\sum_{i=1}^{n} \mathbb{E}(X_i^2) - n\,\mathbb{E}(\bar{X}^2)\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} (\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right] = \sigma^2.
\end{aligned}$$

此时，我们称样本方差 S^2 是总体方差 σ^2 的一个无偏估计器 (unbiased estimator，一个随机变量)，其观察值 s^2 称为无偏估计 (unbiased estimate，一个数字)。可见系数 $\frac{1}{n-1}$ 正是为了达到“无偏”的效果。

进一步假设总体 $X \sim N(\mu, \sigma^2)$。那么由书本 106 页 (2.8) 式知 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 也服从正态分布，具体表述为:

定理

$$\bar{X} \sim N(\mu, \sigma^2/n).$$

我们还有以下两个重要定理。

定理 (3)

1. $$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$
2. $\bar{X}$ 和 S^2 相互独立。

证明见本课件末尾。

定理 (4)

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1).$$

证明.

由前面两个定理知

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1), \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$$

且两者相互独立。由 t 分布的定义可得

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \Bigg/ \sqrt{\frac{S^2}{\sigma^2}} \sim t(n-1).$$

化简即得到定理结论。 □

对于两个正态总体，我们有如下定理。

定理 (5)

设 $X_1, \ldots, X_{n_1}$ 和 $Y_1, \ldots, Y_{n_2}$ 分别是来自正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的样本，且这两组样本相互独立[a]。设 $\bar{X}, \bar{Y}$ 为对应的样本均值，S_1^2, S_2^2 为样本方差。那么有

1.
$$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1).$$

2. 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时，
$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_W\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2),$$

   其中
$$S_W^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}, \quad S_W = \sqrt{S_W^2}.$$

[a]意思是随机向量 $[X_1, \ldots, X_{n_1}]$ 和 $[Y_1, \ldots, Y_{n_2}]$ 相互独立。

证明.

1. 由定理 (3) 知

$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1), \quad \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2-1). \tag{2}$$

由假设知 S_1^2, S_2^2 相互独立，所以由 F 分布的定义知

$$\frac{S_1^2}{\sigma_1^2} \Big/ \frac{S_2^2}{\sigma_2^2} \sim F(n_1-1, n_2-1).$$

2. 由两组样本的独立性知 $\bar{X}-\bar{Y} \sim N(\mu_1-\mu_2, \frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2})$，因此

$$U := \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim N(0,1).$$

□

证明（续）.

又由条件 (2)、独立性、χ^2 分布的可加性知

$$V := \frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_1+n_2-2).$$

由定理 (3) 的证明及其推广可知 U, V 相互独立。因此，按照 t 分布的定义，我们有

$$\frac{U}{\sqrt{V/(n_1+n_2-2)}} = \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_W\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2).$$

□

定理 (3) 证明简述.

- 做归一化。令 $Z_i = \frac{X_i-\mu}{\sigma}$，$i = 1, \ldots, n$。则由定理假设知 $Z_1, \ldots, Z_n$ 是服从 $N(0, 1)$ 的 i.i.d. 随机变量序列，且

$$\bar{Z} = \frac{\bar{X}-\mu}{\sigma}, \quad \frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n} Z_i^2 - n\bar{Z}^2.$$

- 取一 $n \times n$ 正交矩阵 A，其中第一行元素均为 $1/\sqrt{n}$。做正交（线性）变换

$$Y = AZ, \text{ 其中 } Y = [Y_1, \cdots, Y_n]^\top, \; Z = [Z_1, \cdots, Z_n]^\top.$$

易知 $Y_1, \ldots, Y_n$ 仍为正态随机变量。进一步由于 $Z_i \sim N(0, 1)$，故

$$\mathbb{E}(Y) = \mathbf{0}, \quad \mathrm{Var}(Y) := \mathbb{E}(YY^\top) = I_n,$$

即 $Y_1, \ldots, Y_n$ 两两不相关，由正态随机变量的性质可知它们相互独立，且 $Y_i \sim N(0, 1)$，$i = 1, \ldots, n$。

□

定理 (3) 证明简述（续）.

- 计算 $Y_1 = \sqrt{n}\bar{Z}$, $\sum_{i=1}^n Y_i^2 = Y^\top Y = \sum_{i=1}^n Z_i^2$，故

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n Z_i^2 - n\bar{Z}^2 = \sum_{i=2}^n Y_i^2 \sim \chi^2(n-1).$$

  此即定理第一部分结论。

- 由于

$$\bar{X} = \sigma\bar{Z} + \mu = \frac{\sigma}{\sqrt{n}} Y_1 + \mu$$

  仅依赖 Y_1，而 $S^2 = \frac{\sigma^2}{n-1}\sum_{i=2}^n Y_i^2$ 仅依赖于 $Y_2, \ldots, Y_n$。由 $Y_1, Y_2, \ldots, Y_n$ 的独立性知 $\bar{X}$ 和 S^2 相互独立。定理第二部分证毕。

$\square$

定理 (3 向多个同方差正态总体的推广)

设 $\bar{X}, \bar{Y}, S_1^2, S_2^2$ 是定理 (5) 第二点中所说的正态总体 $N(\mu_1, \sigma^2)$，$N(\mu_2, \sigma^2)$ 的样本均值和样本方差，则 $\bar{X}, \bar{Y}, S_1^2, S_2^2$ 相互独立。

对于 $m \geq 2$ 个同方差正态总体，设 $\bar{X}_i, S_i^2$ 分别是总体 $N(\mu_i, \sigma^2)$ 的样本均值和样本方差，$i = 1, \ldots, m$，且设各样本相互独立，则 $\bar{X}_1, \ldots, \bar{X}_m, S_1^2, \ldots S_m^2$ 相互独立。

The End