Ch 6. 数理统计的基本概念

钟友良

zhongyl@scut.edu.cn

Outline

6.0 前言

6.1 总体、样本和统计量

6.3 抽样分布

Outline

Outline

概率与统计

概率论和统计的区别:

- 概率论: 已知随机变量 X 的分布 (分布列或者分布密度函数), 进而根据分布进行事件概率和矩计算
- 统计: 总体的分布未知或者不完全已知的情形下, 通过抽样并研究抽样得到数据的统计特征, 并估计总体的分布。

问题提出

对某个学校的小学生身高 X 进行调查，确定其分布函数？X 为待研究的 总体 ，构成总体的每个成员被称为一个 个体 。

研究方法

对总体进行 简单随机抽样 ， 并根据抽样结果推断整体。

抽样：从 X 中随机的抽出 n 个个体，记为 n 维样本 。由于抽取的个体是随机的，即无法提前预知其数值，样本为 n 维随机向量 $(X_1, X_2, \cdots, X_n)$。

简单随机抽样

简单随机抽样:

- 样本与总体 同分布
- $X_1, X_2, \cdots, X_n$ 相互独立

总体与样本

- 总体: 总体是一个分布待确定的随机变量 X。总体中包含很多个 个体 。
- 样本: 从总体中随机抽取的 n 个个体，为 n 维随机向量 $(X_1, X_2, \cdots, X_n)$。对样本进行一次观测，记 $(x_1, x_2, \cdots, x_n)$ 为样本的一个 观测值 。

分布函数

命题 6.1.1:
设总体 $X \sim F_X(\cdot, \theta)$, $(X_1, X_2, \cdots, X_n)$ 为其样本,
那么

$$F_{X_1,X_2,\cdots,X_n}(x_1, x_2, \cdots, x_n; \theta)$$
$$= F_X(x_1, \theta) F_X(x_2, \theta) \cdots F_X(x_n, \theta)$$

经验分布函数

问题：由样本估计总体的分布函数。

经验分布函数：设总体的样本 $(X_1, X_2, \cdots, X_n)$ 的一次观测值为 $(x_1, x_2, \cdots, x_n)$，并将其进行升序排列 $x_{(1)} \le x_{(2)} \le \cdots \le x_{(10)}$，并定义

$$F_n^X(x) = \begin{cases} 0, & x < x_{(1)} \\ \cdots & \\ \frac{k}{n}, & x \in [x_{(k)}, x_{(k+1)}) \\ \cdots & \\ 1, & x \ge x_{(n)} \end{cases}$$

称 F_n^X 为总体 X 的一个经验分布函数，或样本分布函数。

经验分布函数的例子

- 抛硬币
- 投骰子

Glivenko 定理

设 $X_1, X_2, \cdots, X_n$ 是取自总体 X 的容量为 n 的样本，总体的分布函数为 $F_X(x)$，$F_n(x)$ 为其经验分布函数，那么

$$P\left(\lim_{n\to\infty} sup_{-\infty<x<\infty} F_n^X(x) - F_X(x) = 0\right) = 1$$

上面的结果说明，就是当样本容量足够大的时候，用经验分布函数来推断总体的分布函数是合理的。

练习

某食品厂生产饮料，现在从生产线 (X) 上任意抽取 5 瓶，称得重量依次为 (g)

$$351 \quad 347 \quad 351 \quad 344 \quad 351$$

请写出总体 X 的经验分布函数

统计量

注意下面这个关系

$$\text{样本} \xrightarrow{\text{统计量}} \text{总体}$$

统计量是通过样本研究总体的手段
统计量是样本的函数，不含有任何未知参数，即统计量由样本唯一决定。

常用统计量

- 样本均值:

$$\overline{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- 样本方差:

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

常用统计量

- 修正样本方差

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

- 样本的 k 阶原点矩

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

- 样本的 k 阶中心矩

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

常用统计量

- 顺序统计量 $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$，其中 $X_{(1)} = \min\{X_1, X_2, \cdots, X_n\}$，$X_{(n)} = \max\{X_1, X_2, \cdots, X_n\}$，而 $X_{(k)}$ 是将 $X_1, X_2, \cdots, X_n$ 的取值从小到大排列的第 k 位的值。
- 样本中位数

$$\tilde{X} = \begin{cases} X_{\left(\frac{n+1}{2}\right)}, & \text{若}n\text{为奇数} \\ \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right), & \text{若}n\text{为偶数} \end{cases}$$

- 样本极差

$$R_n^X = X_{(n)} - X_{(1)}$$

Outline

抽样分布

设计合适的统计量可以用来研究总体的数字特征和参数。

命题 6.3.1: 设 $(X_1, X_2, \cdots, X_n)$ 是来自总体 X 的样本，$E[X] = \mu$，$Var[X] = \sigma^2$，那么

- $$E[\overline{X}] = \mu, \quad Var[\overline{X}] = \frac{\sigma^2}{n}$$

- $$E[S_n^2] = \frac{n-1}{n}\sigma^2, \quad E[S_n^{*2}] = \sigma^2$$

抽样分布

证明:

$$E[\overline{X}] = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

再由 X_i 之间的独立性有

$$Var[\overline{X}] = Var[\frac{1}{n}\sum_{i=1}^{n} X_i] = \frac{1}{n^2}\sum_{i=1}^{n} Var[X_i]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}$$

抽样分布

对于 S_n

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

$$= \frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 - 2\overline{X}\sum_{i=1}^{n}X_i + n\overline{X}^2\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \overline{X}^2$$

抽样分布

所以

$$
\begin{aligned}
E[S_n^2] &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i^2] - E[\overline{X}^2] \\
&= \frac{1}{n}\sum_{i=1}^{n}(Var[X_i] + (E[X_i])^2) - (Var[\overline{X}] + (E[\overline{X}])^2) \\
&= \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2 \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}
$$

$$
E[S_n^{*2}] = E\left[\frac{n}{n-1}S_n^2\right] = \frac{n}{n-1}E[S_n^2] = \sigma^2
$$

χ^2 分布

命题 6.3.2:
设总体 $X \sim N(0,1)$，$(X_1, X_2, \cdots, X_n)$ 为其简单随机样本，则

$$\boldsymbol{X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)}$$

其中

$$\chi^2(n) = \Gamma(\frac{n}{2}, \frac{1}{2}).$$

χ^2 分布

若 $X \sim \Gamma(\frac{n}{2}, \frac{1}{2})$，则称 X 服从自由度为 n 的 χ^2 分布，记为 $X \sim \chi^2(n)$，其分布密度函数为

$$f_X(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & \text{其他} \end{cases}$$

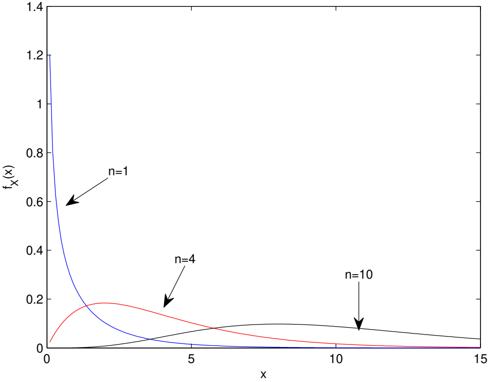
$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

χ^2 分布图像

根据例题 2.4.2，当 $X \sim N(0,1)$，

$$X^2 \sim \chi^2(1)$$

即 X^2 服从自由度为 1 的 χ^2 分布。



χ^2 分布的性质

- 若 $X \sim \chi^2(n)$，那么 $E[X] = n$， $Var[X] = 2n$。
- 若 $X_1 \sim \chi^2(n_1)$， $X_2 \sim \chi^2(n_2)$， 且 X_1 与 X_2 独立，那么 $X_1 + X_2 \sim \chi^2(n_1 + n_2)$。
- 若 $X \sim \chi^2(n)$，那么 $\frac{X-n}{\sqrt{2n}}$ 近似服从 $N(0,1)$，当 $n \to \infty$

习题 6.2

设 $(X_1, X_2, \cdots, X_5)$ 是来自总体 $N(0, 4)$ 的一个样本，且

$$Y = aX_1^2 + b(2X_2 + 3X_3)^2 + c(4X_4 - X_5)^2$$

请问 $a, b, c > 0$ 取什么值的时候，随机变量 Y 服从 $\chi^2(n)$ 分布，n 为多少?

习题 6.6

设 $X_1, X_2, \cdots, X_{20}$ 是来自总体 $X \sim N(0,1)$ 的一个样本，记

$$Y = \frac{1}{10}\left(\sum_{i=1}^{10} X_i\right)^2 + \frac{1}{10}\left(\sum_{i=11}^{20} X_i\right)^2$$

请问 Y 服从什么分布?

t 分布

命题 6.3.3: 设 $X \sim N(0,1)$, $Y \sim \chi^2(n)$, 且 X 与 Y 相互独立, 令

$$\boldsymbol{T = \frac{X}{\sqrt{Y/n}} \sim t(n)}$$

则 T 的分布密度函数 f_T 为

$$f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma(\frac{n}{2})}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

t 分布图像

$X \sim t(n)$ 表示 X 服从自由度为 n 的 t 分布

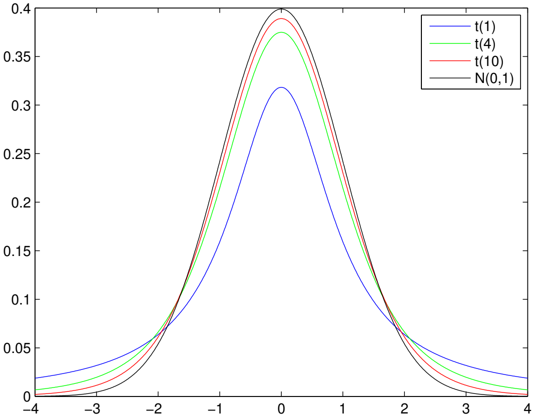


图: $t(n)$ 分布密度函数

t 分布的性质

- t 分布关于原点对称
- 当 $n \to \infty$，$t(n)$ 趋于标准正态分布
- $$E[X] = 0(n > 1), Var[X] = \frac{n}{n-2}(n > 2)$$

习题 6.4

设 $(X_1, X_2, \cdots, X_9)$ 是来自总体 $N(0,1)$ 的简单随机样本，请确定正数 C，使得

$$\frac{C(X_1+X_2+X_3)}{\sqrt{(X_4+X_5)^2+(X_6+X_7)^2+(X_8+X_9)^2}}$$

服从 t 分布，并指出其自由度。

F 分布

命题 6.3.4：若 $X \sim \chi^2(m)$，$Y \sim \chi^2(n)$，且 X 与 Y 相互独立，令

$$\boldsymbol{Z = \frac{X/m}{Y/n} \sim F(m, n)},$$

则 Z 的分布密度函数为

$$f_Z(x) = \begin{cases} \dfrac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} m^{\frac{n}{2}} n^{\frac{m}{2}} \dfrac{x^{m/2-1}}{(mx+n)^{(m+n)/2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

称随机变量 Z 服从第一自由度为 m，第二自由度为 n 的 F 分布，记作 $Z \sim F(m, n)$。

F 分布图像

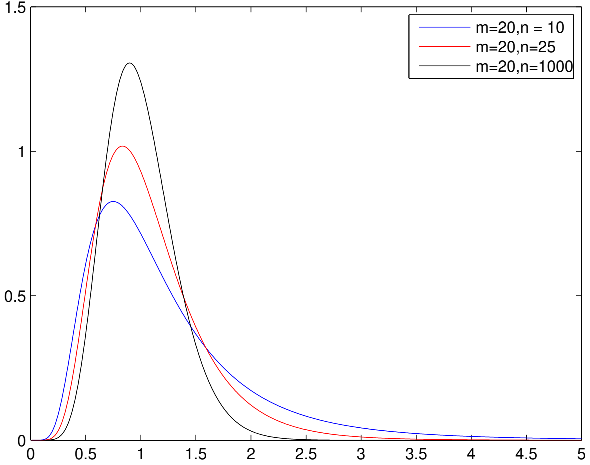


图: $F(m, n)$ 分布密度函数

F 分布

$Z \sim F(m,n)$，那么 $\frac{1}{Z} \sim \boldsymbol{F(n,m)}$。
证明：

$$
\begin{aligned}
F_{\frac{1}{Z}}(t) &= P(\frac{1}{Z} \le t) \\
&= P(Z \ge \frac{1}{t}) \\
&= 1 - P(Z < \frac{1}{t}) \\
&= 1 - \int_0^{\frac{1}{t}} f_Z(x)\,dx
\end{aligned}
$$

F 分布

$$
\begin{aligned}
f_{\frac{1}{Z}}(t) = \frac{d}{dt} F_{\frac{1}{Z}}(t) &= t^{-2} f_Z(\frac{1}{t}) \\
&= t^{-2} C_{m,n} \frac{t^{1-\frac{m}{2}}}{\left(mt^{-1}+n\right)^{\frac{m+n}{2}}} \\
&= t^{-2} C_{m,n} \frac{t^{\frac{m+n}{2}-\frac{m}{2}+1}}{\left(nt+m\right)^{\frac{m+n}{2}}} \\
&= C_{m,n} \frac{t^{\frac{n}{2}-1}}{\left(nt+m\right)^{\frac{m+n}{2}}}
\end{aligned}
$$

其中

$$
C_{m,n} = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} m^{\frac{n}{2}} n^{\frac{m}{2}}
$$

习题 6.3

设 $X_1, X_2, \cdots, X_8$ 是来自总体 $N(0,1)$ 的简单随机样本，求常数 c，使得

$$\frac{c(X_1^2+X_2^2)}{(X_3+X_4+X_5)^2+(X_6+X_7+X_8)^2}$$

服从 F 分布，并指出其自由度

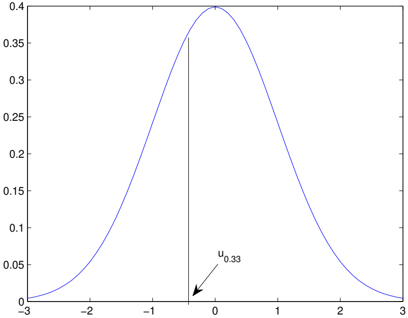
分位数

分位数: 设 $X \sim \phi(n)$，$0 < \alpha < 1$，称满足

$$P(X \leq \phi_\alpha(n)) = \alpha$$

的实数 $\phi_\alpha(n)$ 为分布 $\phi(n)$ 的 α 分位数，或叫做 α 分位点。

分位数



上图是标准正态分布的密度区线，其中 $u_{0.33}$ 是其 0.33 分位点，表示

$$P(X \leq u_{0.33}) = 0.33$$

计算 t 分位数

若 $X \sim t(n)$，由于 $\lim_{n\to\infty} X \sim N(0,1)$，

$$\lim_{n\to\infty} t_\alpha(n) = u_\alpha$$

练习

$t_{0.95}(10), t_{0.1}(20), t_{0.9}(50)$

计算 χ 分位数

- 若 $X \sim \chi^2(n)$ 分布，当 $n \to \infty$，$(X-n)/\sqrt{2n} \sim N(0,1)$，所以

$$\begin{aligned}\alpha &= P(X \le \chi^2_\alpha(n)) \\ &= P((X-n)/\sqrt{2n} \le (\chi^2_\alpha(n)-n)/\sqrt{2n})\end{aligned}$$

即

$$u_\alpha = (\chi^2_\alpha(n)-n)/\sqrt{2n}$$

解出

$$\boldsymbol{\chi^2_\alpha(n) = n + \sqrt{2n}u_\alpha}, \quad n \to \infty$$

练习

$\chi^2_{0.99}(10), \chi^2_{0.05}(20), \chi^2_{0.95}(60)$

计算 F 分位数

若 X 服从 $F(m,n)$，则有 $\frac{1}{X} \sim F(n,m)$，若已知 $F_\alpha(m,n)$，那么

$$\alpha = P(X < F_\alpha(m,n)) = P(\frac{1}{X} > \frac{1}{F_\alpha(m,n)})$$

$$= 1 - P(\frac{1}{X} \le \frac{1}{F_\alpha(m,n)})$$

即

$$P(\frac{1}{X} \le \frac{1}{F_\alpha(m,n)}) = 1 - \alpha$$

也即

$$\boldsymbol{\frac{1}{F_\alpha(m,n)} = F_{1-\alpha}(n,m)}$$

练习

$F_{0.99}(5,4), F_{0.05}(3,7)$

习题 5.8

一个系统由 n 个独立的部件组成，每个部件损坏的概率为 0.1。系统的正常运行需要 $\geq 80\%$ 的部件正常工作。
求：n 需要多少才能使得系统正常运行的概率 $\geq 95\%$?

$X_i \sim B(1,p)$. $p=0.9$

$$\frac{\sum X_i - E[\sum X_i]}{\sqrt{Var[\sum X_i]}} \longrightarrow$$

$E[\sum X_i] = np$, $Var[\sum X_i] = np(1-p)$

$$P\left(\sum_{i=1}^{n} X_i \geq 0.8n\right) \geq 0.95.$$

$$= 1 - P\left(\sum X_i < 0.8n\right)$$

$$= 1 - P\left(\frac{\sum X_i - E[\sum X_i]}{\sqrt{Var[\sum X_i]}} < \frac{0.8n - np}{\sqrt{np(1-p)}}\right)$$

$$\frac{0.8n - 0.9n}{\sqrt{n \cdot 0.9 \times 0.1}} = -\frac{\sqrt{n}}{3}$$

$$\approx 1 - P\left(Y < -\frac{\sqrt{n}}{3}\right) = P\left(Y < \frac{\sqrt{n}}{3}\right) = \Phi\left(\frac{\sqrt{n}}{3}\right)$$

$Y \sim N(0,1)$

$$\Phi\left(\frac{\sqrt{n}}{3}\right) \geq \Phi(u_{0.95})$$

$$\Rightarrow \frac{\sqrt{n}}{3} \geq u_{0.95}$$

查表 $\Phi(1.64) \leq 0.95 \leq \Phi(1.65)$

回顾：1 阶统计量

$$X \sim N(0,1) \underset{X=\sigma^{-1}(Y-\mu)}{\overset{Y=\sigma X+\mu}{\rightleftarrows}} Y \sim N(\mu, \sigma^2)$$

$\bar{X} = \frac{1}{n}\sum X_i \sim N(0, \frac{1}{n})$.

$(\sqrt{n}\,\bar{X} \sim N(0,1))$

$\bar{Y} = \sigma\bar{X} + \mu \sim N(\mu, \frac{\sigma^2}{n})$

$(\sigma^{-1} \cdot \sqrt{n}\,\bar{Y} \sim N(0,1))$

$X_i - \bar{X} \sim N(0, \frac{n-1}{n})$

$(\sqrt{\frac{n}{n-1}}(X_i - \bar{X}) \sim N(0,1))$

$Y_i - \bar{Y} = \sigma(X_i - \bar{X}) \sim N(0, \frac{n-1}{n}\sigma^2)$

$(\sigma \cdot \sqrt{\frac{n}{n-1}}(Y_i - \bar{Y}) \sim N(0,1))$

$\sum \left(\sqrt{\frac{n}{n-1}}(X_i - \bar{X})\right)^2 \sim \chi^2(n)$.

$\parallel$

$\frac{n}{n-1} \cdot \sum (X_i - \bar{X})^2$ 对吗？

Cochran’s Theorem 的应用

By Cochran’s Theorem, $X^T AX \sim \chi^2(\text{rank}(A))$
where $X \sim N(0,1)$
$\sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2(n-1)$ where $X_i \sim N(0,1)$.

$$\sum_i (X_i - \bar{X})^2 = \sum_i \Big(X_i - \frac{1}{n}\sum_j X_j\Big)^2$$
$$= \sum_i \Big(X_i^2 - \frac{2}{n} X_i \Big(\sum_j X_j\Big) + \frac{1}{n^2}\Big(\sum_j X_j\Big)^2\Big)$$
$$= \sum_i X_i^2 - \frac{2}{n}\Big(\sum X_i\Big)\cdot\Big(\sum_j X_j\Big) + \frac{1}{n}\Big(\sum_j X_j\Big)^2$$
$$= \sum X_i^2 - \frac{1}{n}\Big(\sum X_i\Big)^2$$
$$= \sum X_i^2 - \frac{1}{n}\Big(\sum X_i^2 + \sum_{i \neq j} X_i X_j\Big).$$
$$= (X_1 \cdots X_n)\underbrace{\left(\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} - \frac{1}{n}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}\right)}_{I_n - \frac{1}{n}\cdot A_1}\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

Claim: $\text{rank}(I_n - \frac{1}{n}A_1) = n-1$.

Pf: ① $rk \geq \underbrace{rk(I_n)}_{= n} - \underbrace{rk(\frac{1}{n}A_1)}_{= 1} = n-1$

② $(I_n - \frac{1}{n}A_1)\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

$\Rightarrow rk < n$

①+② $\Rightarrow rk = n-1$.

. . .

回顾：2 阶统计量

$$X \sim N(0,1) \xrightarrow{Y=\sigma X+\mu} Y \sim N(\mu, \sigma^2)$$

$$X = \frac{Y-\mu}{\sigma}.$$

$X^2 \sim \chi^2(1).$

$\sum X_i^2 \sim \chi^2(n).$

$\sum (X_i - \bar{X})^2 \sim \chi^2(n-1)$

$\quad = n \cdot S_n^2(X)$

$= (n-1) S_n^{*2}(X).$

$\left(\frac{Y-\mu}{\sigma}\right)^2 \sim \chi^2(1)$

$\sigma^{-2} \sum (Y_i - \bar{Y})^2 = \sum (X_i - \bar{X})^2 \sim \chi^2(n-1)$

$= \sigma^{-2} \cdot n \cdot S_n^2(Y)$

$= \sigma^{-2} (n-1) \cdot S_n^{*2}(Y).$

正态总体的抽样分布

抽样分布基本定理：设总体 $X \sim N(0,1)$，$(X_1, X_2, \cdots, X_n)$ 为其样本，则样本均值 $\boldsymbol{\overline{X} \sim N(0, \frac{1}{n})}$，$\boldsymbol{nS_n^2 \sim \chi^2(n-1)}$，并且 $\overline{X}$ 与 S_n^2 相互独立。

- 对于服从正态分布的总体，样本均值 $\overline{X}$ 服从一维正态分布。
- $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$，即 S_n^2 是 $\overline{X}$ 的函数，但是它们却相互独立。

推论 6.3.1:

设总体 $X \sim N(\mu, \sigma^2)$, $(X_1, X_2, \cdots, X_n)$ 为其样本，则 $\overline{X}$ 与 S_n^2 相互独立，且

$$\boldsymbol{\overline{X} \sim N(\mu, \sigma^2/n)}$$
$$\boldsymbol{\frac{(n-1)S_n^{*2}}{\sigma^2} = \frac{nS_n^2}{\sigma^2} \sim \chi^2(n-1)}$$

推论 6.3.2: (ex 6.5).

设总体 $X \sim N(\mu, \sigma^2)$，$(X_1, X_2, \cdots, X_n)$ 为其样本，那么

$$\frac{\overline{X}-\mu}{S_n^*}\sqrt{n} = \frac{\overline{X}-\mu}{S_n}\sqrt{n-1} \sim t(n-1)$$

推论 6.3.3: (ex 6.7)

设总体 $X \sim N(\mu_1, \sigma_1^2)$，$(X_1, X_2, \cdots, X_m)$ 为其样本，样本均值为 $\overline{X}$，样本方差为 S_{1m}^2；另有与 X 独立的总体 $Y \sim N(\mu_2, \sigma_2^2)$，$(Y_1, Y_2, \cdots, Y_n)$ 为其样本，样本均值为 $\overline{Y}$，样本方差为 S_{2n}^2，那么

$$\frac{mS_{1m}^2}{nS_{2n}^2}\frac{\sigma_2^2}{\sigma_1^2}\frac{n-1}{m-1} = \frac{S_{1m}^{*2}}{S_{2n}^{*2}}\frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1, n-1)$$

推论 6.3.4:

在推论 6.3.3 的假定中增加一个条件 $\sigma_1 = \sigma_2$,
那么

$$\frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{mS_{1m}^2 + nS_{2n}^2}{m+n-2}}\sqrt{\frac{1}{m} + \frac{1}{n}}} \sim t(m+n-2)$$

命题 6.3.5:

设总体 X 的分布函数为 F_X，分布密度函数为 f_X，那么

$$f_{X_{(k)}}(x) = \frac{n!}{(n-k)!(k-1)!}[F_X(x)]^{k-1}[1-F_X(x)]^{n-k}$$

$k = 1, 2, \cdots, n$

习题 6.5

设 $X_1, X_2, \cdots, X_{10}$ 是来自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本，且

$$\overline{X} = \frac{1}{9}\sum_{i=1}^{9} X_i,$$

$$S_9^{*2} = \frac{1}{8}\sum_{i=1}^{9}(X_i - \overline{X})^2,$$

$$T = \frac{3(X_{10} - \overline{X})}{S_9^{*}\sqrt{10}}$$

请问 T 服从何种分布

习题 6.7

设总体 $X \sim N(0, \sigma^2)$，从 X 中抽取样本 $(X_1, X_2, \cdots, X_{14})$，且

$$Y_1 = \frac{1}{5}\sum_{i=1}^{5} X_i,\ Y_2 = \frac{1}{5}\sum_{i=10}^{14} X_i,$$

$$Z_1 = \sum_{i=1}^{5}(X_i - Y_1)^2,\ Z_2 = \sum_{i=10}^{14}(X_i - Y_2)^2,$$

$$Z_3 = \sum_{i=6}^{9} X_i^2,\ T = \frac{Z_1 + Z_2}{2Z_3}$$

请问 T 服从什么分布