1. $\xi_1, \xi_2, \xi_3, \xi_4$ 独立，$\xi_i \sim N(0,1)$ $i=1,2,3,4$

求 $\dfrac{\sqrt{3}\xi_1}{\sqrt{\xi_2^2+\xi_3^2+\xi_4^2}}$ 的分布。

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

解： $\xi_1 \sim N(0,1)$ $\xi_2^2+\xi_3^2+\xi_4^2 \sim \chi^2(3)$

且两者相互独立，

$$\frac{\xi_1}{\sqrt{\dfrac{\xi_2^2+\xi_3^2+\xi_4^2}{3}}} = \frac{\sqrt{3}\xi_1}{\sqrt{\xi_2^2+\xi_3^2+\xi_4^2}} \sim t(3)$$

2. $X_1, X_2, \cdots, X_{15}$ 独立，均服从 $N(0,2^2)$

$$Y=\frac{X_1^2+X_2^2+\cdots+X_{10}^2}{2(X_{11}^2+X_{12}^2+\cdots+X_{15}^2)}$$ 求 Y 的分布。

解：$\frac{1}{4}(X_1^2+X_2^2+\cdots+X_{10}^2) \sim \chi^2(10)$

$\frac{1}{4}(X_{11}^2+X_{12}^2+\cdots+X_{15}^2) \sim \chi^2(5)$

$$F=\frac{X/n_1}{Y/n_2}$$

且两者相互独立，

$$Y=\frac{(X_1^2+X_2^2+\cdots+X_{10}^2)/10}{(X_{11}^2+X_{12}^2+\cdots+X_{15}^2)/5} \sim F(10,5)$$

一、正态总体的抽样分布

问题一 设定正态总体 $X \sim N(\mu,\sigma^2)$, 并从中抽出一组样本$(X_1, X_2, \cdots, X_n)$, 请确定如下统计量的分布:

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

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

定理1

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

3. $\bar{X}$ 与 S^2 相互独立.

定理

设定正态总体 $X \sim N(\mu,\sigma^2)$, 并从中抽出一组样本 $(X_1, X_2, \cdots, X_n)$, 则

1. $U = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$;

2. $\dfrac{(n-1)}{\sigma^2} S^2 \sim \chi^2(n-1)$;

3. $\overline{X}$ 与 S^2 相互独立.

定义

仅含总体分布中一个未知参数且分布已知的样本函数 $g(X_1, X_2, \cdots, X_n)$ 称为**枢轴量**。

例如，$\dfrac{(n-1)}{\sigma^2}S^2$ 是关于总体方差σ^2的枢轴量，但是

$U=\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ 在 σ 和 μ 均未知时不是枢轴量。

这时，通常用样本标准差 S 替换 U 中未知的总体标准差 σ, 得到

$$T=\frac{\overline{X}-\mu}{S/\sqrt{n}}.$$

问题二

设定正态总体 $X\sim N(\mu,\sigma^2)$, 并从中抽出一组样本 $(X_1,X_2,\cdots,X_n)$, 请确定 T 的分布.

问题二

设定正态总体 $X \sim N(\mu,\sigma^2)$, 并从中抽出一组样本 $(X_1, X_2, \cdots, X_n)$, 请确定 T 的分布，其中

$$T = \frac{\overline{X}-\mu}{S/\sqrt{n}}.$$

$$T = \frac{U}{\sqrt{\frac{(n-1)S^2}{\sigma^2} \Big/ n-1}}$$

一般化

设 $X \sim N(0,1), Y \sim \chi^2(n)$ 且相互独立, 求

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

的分布

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

定理2 $$\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

证明：$\overline{X}\sim N(\mu,\frac{\sigma^2}{n})$，$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$,

$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，且它们独立。

则由t-分布的定义：$$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\Bigg/\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}\sim t(n-1)$$

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

定理2

设定正态总体 $X \sim N(\mu,\sigma^2)$, 并从中抽出一组样本 $(X_1, X_2, \cdots, X_n)$, 则

1. $U = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$; 2. $\dfrac{(n-1)}{\sigma^2}S^2 \sim \chi^2(n-1)$;

2. $T = \dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$.

问题三

设 $X \sim N(\mu_1,\sigma_1^2)$ 和 $Y \sim N(\mu_2,\sigma_2^2)$ 是两个相互独立正态总体且各抽得样本 $(X_1,\cdots,X_{n_1})$ 和 $(Y_1,\cdots,Y_{n_2})$,求样本均值差 $(\overline{X}-\overline{Y})$的分布.

定理3

1. $U=\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \sim N(0,1)$;

当$\sigma_1=\sigma_2$时,令$S_{12}^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$,则

2. $T=\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_{12}\sqrt{1/n_1+1/n_2}} \sim t(n_1+n_2-2)$.

定理3 设 $X_1, X_2, \cdots, X_{n_1}$ 与 $Y_1, Y_2, \cdots Y_{n_2}$ 分别是具有相同方差的两个正态总体$N(\mu_1, \sigma^2), N(\mu_2, \sigma^2)$的样本，且它们独立.

设$\overline{X} = \dfrac{1}{n_1}\sum_{i=1}^{n_1} X_i$，$\overline{Y} = \dfrac{1}{n_2}\sum_{j=1}^{n_2} Y_j$ 分别是两个样本的均值。

$$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_i - \overline{X})^2, \quad S_2^2 = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}(Y_j - \overline{Y})^2$$

分别是两个样本的方差. 则有:

$$\frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

证明: $\overline{X}-\overline{Y}\sim N(\mu_1-\mu_2,\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}),$

所以 $$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{1/n_1+1/n_2}}\sim N(0,1),$$

且 $\dfrac{(n_1-1)S_1^2}{\sigma^2}\sim\chi^2(n_1-1),\quad \dfrac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_2-1),$

它们独立.

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

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

由$t-$分布的定义：

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{1/n_1+1/n_2}}\Bigg/\sqrt{\frac{(n_1-1)S_1^2}{\sigma^2}+\frac{(n_2-1)S_2^2}{\sigma^2}/(n_1+n_2-2)}$$

$$\sim t(n_1+n_2-2)$$

即：$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$$

问题四

设 $X \sim N(\mu_1,\sigma_1^2)$ 和 $Y \sim N(\mu_2,\sigma_2^2)$ 是两个相互独立正态总体，且各抽得样本$(X_1,L\ ,X_{n_1})$ 和 $(Y_1,L\ ,Y_{n_2})$.求样本方差比 $\dfrac{S_1^2}{S_2^2}$ 的分布.

$$F=\frac{\dfrac{(n_1-1)S_1^2\big/\sigma_1^2}{n_1-1}}{\dfrac{(n_2-1)S_2^2\big/\sigma_2^2}{n_2-1}}$$

一般化

设 $X \sim \chi^2(m), Y \sim \chi^2(n)$ 且相互独立，求

$$Z=\frac{X/m}{Y/n}$$

的分布.

定理4

设 $X \sim N(\mu_1, \sigma_1^2)$ 和 $Y \sim N(\mu_2, \sigma_2^2)$ 是两个相互独立正态总体，且各抽得样本$(X_1, \cdots, X_{n_1})$ 和 $(Y_1, \cdots, Y_{n_2})$.求样本方差比 $\dfrac{S_1^2}{S_2^2}$ 的分布.

$$\mathrm{F} = \left(\frac{\sigma_2}{\sigma_1}\right)^2 \frac{\mathrm{S}_1^2}{\mathrm{S}_1^2} \sim \mathrm{F}(n_1 - 1, n_2 - 1)$$

定理4 设 $X_1, X_2, \cdots, X_{n_1}$ 与 $Y_1, Y_2, \cdots Y_{n_2}$ 分别是两个正态总体 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ 的样本，且它们独立。

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

$$F = \frac{X/n_1}{Y/n_2}$$

证明：

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

它们独立.

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

小结：什么是数理统计

数理统计学以简单随机样本为依据，利用统计量或枢轴量的抽样分布，对所研究的问题的总体分布做统计推断。

例10 设X_1, X_2, X_3是总体$N(2,9)$的样本,

求(1)$P\{\overline{X} > 3\}$; (2)$P\{|\overline{X} - 2| > 1\}$; (3)$P\{S^2 > 26.955\}$;

(4)$P\{\max(X_1, X_2, X_3) > 4\}$; (5)$P\{\min(X_1, X_2, X_3) < 0\}$.

解:(1) 由于 $\overline{X} \sim N(2,3)$,

所以 $P\{\overline{X} > 3\} = 1 - \Phi(\frac{3-2}{\sqrt{3}}) = 1 - \Phi(\frac{1}{\sqrt{3}})$

$= 1 - \Phi(0.58) = 1 - 0.7190 = 0.281$

(2) $P\{|\overline{X} - 2| > 1\} = 1 - P\{|\overline{X} - 2| \le 1\}$

$= 1 - P\{-\frac{1}{\sqrt{3}} \le \frac{\overline{X} - 2}{\sqrt{3}} \le \frac{1}{\sqrt{3}}\}$

$$=1-P\{-\frac{1}{\sqrt{3}}\le\frac{\overline{X}-2}{\sqrt{3}}\le\frac{1}{\sqrt{3}}\}=1-[\Phi(\frac{1}{\sqrt{3}})-\Phi(-\frac{1}{\sqrt{3}})]$$

$$=2-2\Phi(\frac{1}{\sqrt{3}})\ =2\times[1-\Phi(0.58)]$$

$$=2\times[1-0.7190]=0.562$$

$(3)P\{S^2>26.955\};$

（3） 由于 $(3-1)S^2/9\sim\chi^2(2)$，故

$$P\{S^2>26.955\}=P\{2S^2/9>5.99\}\ \approx 0.05$$

(4) $P\{\max(X_1, X_2, X_3) > 4\}$

$$= 1 - P\{\max(X_1, X_2, X_3) \le 4\}$$

$$= 1 - P\{X_1 \le 4, X_2 \le 4, X_3 \le 4\}$$

$$= 1 - P\{X_1 \le 4\}P\{X_2 \le 4\}P\{X_3 \le 4\}$$

$$= 1 - [\Phi(\frac{4-2}{3})]^3$$

$X_1 \sim N(2,9)$

$$= 1 - [\Phi(0.67)]^3$$

$$= 1 - (0.7486)^3$$

$$= 0.58$$

(5) $P\{\min(X_1, X_2, X_3) < 0\}$

$$= 1 - P\{\min(X_1, X_2, X_3) \geq 0\}$$

$$= 1 - P\{X_1 \geq 0, X_2 \geq 0, X_3 \geq 0\}$$

$$= 1 - P\{X_1 \geq 0\}P\{X_2 \geq 0\}P\{X_3 \geq 0\}$$

$$= 1 - [1 - \Phi(\frac{0-2}{3})]^3$$

$X_1 \sim N(2,9)$

$$= 1 - [1 - 1 + \Phi(0.67)]^3$$

$$= 1 - (0.7486)^3$$

$$= 0.58$$

例11 设 $X_1, X_2, \cdots, X_{10}$ 与 $Y_1, Y_2, \cdots Y_{15}$ 分别是正态总体 $N(20,3)$ 的两个独立样本，求 $P\{|\overline{X}-\overline{Y}|>0.1\}$.

解： $\overline{X}-\overline{Y} \sim N(0, \frac{3}{10}+\frac{3}{15})$, 即 $\overline{X}-\overline{Y} \sim N(0,0.5)$.

$$P\{|\overline{X}-\overline{Y}|>0.1\} = 1-P\{|\overline{X}-\overline{Y}| \le 0.1\}$$

$$= 1-P\{\frac{|\overline{X}-\overline{Y}|}{\sqrt{0.5}} \le \frac{0.1}{\sqrt{0.5}}\} = 1-P\{-0.14 \le \frac{\overline{X}-\overline{Y}}{\sqrt{0.5}} \le 0.14\}$$

$$= 2-2\Phi(0.14) \quad = 2-2\times 0.5557$$

$$= 0.8886$$

例12 设$X_1,\cdots,X_n$是总体$X \sim N(\mu,\sigma^2)$的样本,则

$$E\left\{\sum_{i=1}^{n}(X_i-\overline{X})^2\right\}=\sigma^2 E\left\{\sum_{i=1}^{n}(X_i-\overline{X})^2/\sigma^2\right\}$$

$$=\underline{\quad(n-1)\sigma^2\quad}.\qquad (\because \sum_{i=1}^{n}(X_i-\overline{X})^2/\sigma^2 \sim \chi^2(n-1))$$

同理 $$D\left\{\sum_{i=1}^{n}(X_i-\overline{X})^2\right\}=\sigma^4 D\left\{\sum_{i=1}^{n}(X_i-\overline{X})^2/\sigma^2\right\}$$

$$=\underline{\quad 2(n-1)\sigma^4\quad}.$$

例12（续）

$$E\left\{\sum_{i=1}^{n}(X_i-\mu)^2\right\}=\sigma^2 E\left\{\sum_{i=1}^{n}(X_i-\mu)^2/\sigma^2\right\}=\underline{\ n\sigma^2\ }.$$

$$(\because \sum_{i=1}^{n}(X_i-\mu)^2/\sigma^2 \sim \chi^2(n))$$

$$D\left\{\sum_{i=1}^{n}(X_i-\mu)^2\right\}=\sigma^4 D\left\{\sum_{i=1}^{n}(X_i-\mu)^2/\sigma^2\right\}=\underline{\ 2n\sigma^4\ }.$$

$$\frac{(\overline{X}-\mu)\sqrt{n(n-1)}}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}} \sim \underline{\quad t(n-1) \quad};$$

$$\frac{(\overline{X}-\mu)\sqrt{n(n-1)}}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}} = \frac{(\overline{X}-\mu)}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2} \;/\sqrt{n}}$$

$$= \frac{\overline{X}-\mu}{S/\sqrt{n}} \quad \sim t(n-1)$$

例12（续）

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \underline{\chi^2(n)}\,.$$

$$\frac{X_i-\mu}{\sigma} \sim N(0,1)$$

例12（续） $\frac{1}{n}\left(\sum_{i=1}^{n}\frac{X_i-\mu}{\sigma}\right)^2 \sim$ $\underline{\chi^2(1)}$.

因为 $\frac{X_i-\mu}{\sigma} \sim N(0,1), i=1,\cdots,n,$ 且它们独立.

所以 $\sum_{i=1}^{n}\frac{X_i-\mu}{\sigma} \sim N(0,n),$ 故 $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i-\mu}{\sigma} \sim N(0,1),$

则 $\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i-\mu}{\sigma}\right\}^2 \sim \chi^2(1).$

第六章 小 结

1 给出了总体、个体、样本和统计量的概念，要掌握样本均值和样本方差的计算及基本性质。

2 引进了 χ^2 分布、t分布、F分布的定义，会查表计算。

3 掌握正态总体的某些统计量的分布。

作业: 4版147页 1, 2, 3, 4, 9

设$(X_1, X_2, \cdots, X_n, X_{n+1})$ 是取自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本，令

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

试求统计量 $\sqrt{\dfrac{n}{n+1}}\dfrac{X_{n+1}-\bar{X}_n}{S_n}$ 的分布.

设$(X_1, X_2, \cdots, X_n, X_{n+1})$ 是取自总体 $X \sim N(\mu, \sigma^2)$ 的一个样本，令

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

试求统计量$\sqrt{\frac{n}{n+1}}\frac{X_{n+1} - \bar{X}_n}{S_n}$的分布.