数理统计的基本概念

基本概念

总体

$$\text{所研究对象的全体的某项指标, 记为}X$$

样本

$$\begin{aligned}
&\text{总体}X\text{中取}n\text{个个体}X_1, X_2, \cdots, X_n\text{, 称样本}\\
&\text{若①}X_1, \cdots, X_n\text{独立;}\\
&\text{②}X_1, \cdots, X_n\text{与}X\text{同分布.}\\
&\text{称}(X_1, \cdots, X_n)\text{为简单随机样本}\\
&(X_1, \cdots, X_n)\text{的一组取值}(x_1, \cdots, x_n)\text{, 样本观察法}
\end{aligned}$$

统计量

$$\begin{aligned}
&\text{总体}X \Rightarrow (X_1, X_2, \cdots, X_n)\\
&\text{样本的无参函数, 称为统计量}\\
&\text{如}: X \Rightarrow (X_1, X_2, X_3)\\
&\frac{X_1 + X_2 - X_3}{5}\\
&aX_1 + (X_2 - X_3)^2 (\times)\\
&X_1^2 + X_2^2 + X_3^2\\
&X \Rightarrow (X_1, \cdots, X_n)\text{以下为重要的统计量}:\\
&1. \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X} - \text{样本均值}\\
&2. A_k \triangleq \sum_{i=1}^{n} X_i^k (k = 1, 2, \cdots) - \text{样本的}k\text{阶原点矩}\\
&\text{如}A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2\\
&3. S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 - \text{样本方差}
\end{aligned}$$

三个抽样分布

卡方分布

$$\begin{aligned}
&\text{设}X_1, \cdots, X_n\text{相互独立且服从}N(0, 1)\text{分布}\\
&Z = X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)\\
&\text{设}X \sim N(\mu, \sigma^2) \Rightarrow (X_1, \cdots, X_n)\\
&\frac{X_1 - \mu}{\sigma} \sim N(0, 1), \cdots, \frac{X_n - \mu}{\sigma} \sim N(0, 1)\text{且独立}\\
&(\frac{X_1 - \mu}{\sigma})^2 + \cdots + (\frac{X_n - \mu}{\sigma})^2 \sim \chi^2(n)\text{, 即}\\
&\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2 \sim \chi^2(n)
\end{aligned}$$

$$X \sim N(0,4) \Rightarrow (X_1,\cdots,X_4)$$
$$a(X_1+2X_2)^2+b(3X_3-4X_4)^2 \sim \chi^2(n),\text{求}a,b,n.$$

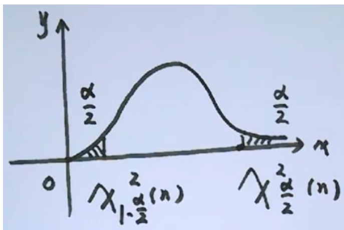
$$X_1+2X_2 \sim N(0,20) \Rightarrow \frac{X_1+2X_2}{\sqrt{20}} \sim N(0,1)$$
$$3X_3-4X_4 \sim N(0,100) \Rightarrow \frac{3X_3-4X_4}{10} \sim N(0,1)\text{且独立}$$
$$\frac{1}{20}(X_1+2X_2)^2+\frac{1}{100}(3X_3-4X_4)^2 \sim \chi^2(2)$$
$$\Rightarrow a=\frac{1}{20},b=\frac{1}{100},n=2$$

性质

$$①X \sim N(0,1) \Rightarrow X^2 \sim \chi^2(1)$$
$$②X \sim \chi^2(m),Y \sim \chi^2(n)\text{且}X,Y\text{独立} \Rightarrow X+Y \sim \chi^2(m+n)$$
$$③X \sim \chi^2(n) \Rightarrow EX=n,DX=2n$$

分位点

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



设总体$X \sim N(0,2)$，(X_1,X_2,X_3)为来自总体X的简单随机样本，$Z=X_1^2+(X_2-X_3)^2$，求$E(Z),D(Z)$.

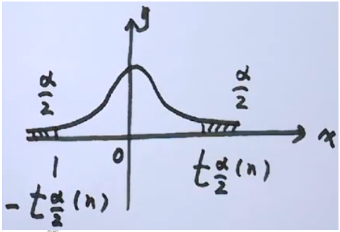
$$X_1 \sim N(0,2) \Rightarrow \frac{X_1}{\sqrt{2}} \sim N(0,1)$$
$$X_2-X_3 \sim N(0,4) \Rightarrow \frac{X_2-X_3}{2} \sim N(0,1)\text{且独立}$$
$$\frac{1}{2}X_1^2 \sim \chi^2(1),\frac{1}{4}(X_2-X_3)^2 \sim \chi^2(1)\text{且独立}$$
$$E(\frac{1}{2}X_1^2)=1 \Rightarrow EX_1^2=2;D(\frac{1}{2}X_1^2)=2 \Rightarrow DX_1^2=8$$
$$E[\frac{1}{4}(X_2-X_3)^2]=1 \Rightarrow E(X_2-X_3)^2=4$$
$$D[\frac{1}{4}(X_2-X_3)^2] \Rightarrow D(X_2-X_3)^2=32$$
$$EZ=2+4=6,DZ=DX_1^2+D(X_2-X_3)^2=8+32=40$$

设总体$X \sim N(0,4)$，(X_1,X_2,X_3)为来自总体X的简单随机样本，
且$a(X_1^2+X_2^2+X_3^2) \sim \chi^2(n)$，求$a,n$.

$$\frac{X_1}{2} \sim N(0,1),\frac{X_2}{2} \sim N(0,1),\frac{X_3}{2} \sim N(0,1)\text{且独立}$$
$$\Rightarrow \frac{1}{4}(X_1^2+X_2^2+X_3^2) \sim \chi^2(3) \Rightarrow a=\frac{1}{4},n=3$$

t分布

$$X \sim N(0,1), Y \sim \chi^2(n)\text{且}X,Y\text{独立,称}$$
$$t = \frac{X}{\sqrt{Y/n}} \sim t(n)$$
$$X \sim t(n) \Rightarrow X \dot{\sim} N(0,1)$$



性质

$$①X \sim t(n) \Rightarrow EX = 0$$
$$②X \sim t(n) \Rightarrow X \dot{\sim} N(0,1), P\{X<0\} = P\{X \geq 0\} = \frac{1}{2}$$

设总体$X \sim N(0,4)$, (X_1,X_2,X_3,X_4)为来自总体X的简单随机样本,求:
(1)$\dfrac{X_1+X_2}{\sqrt{X_3^2+X_4^2}}$服从的分布;(2)$\dfrac{X_1}{|X_2|}$服从的分布.

$$①X_1+X_2 \sim N(0,8) \Rightarrow \frac{X_1+X_2}{2\sqrt{2}} \sim N(0,1)$$
$$\frac{X_3}{2} \sim N(0,1), \frac{X_4}{2} \sim N(0,1)\text{且独立} \Rightarrow \frac{1}{4}(X_3^2+X_4^2) \sim \chi^2(2)\text{且独立}$$
$$\frac{\frac{X_1+X_2}{2\sqrt{2}}}{\sqrt{\frac{1}{4}(X_3^2+X_4^2)/2}} \sim t(2) \Rightarrow \frac{X_1+X_2}{\sqrt{X_3^2+X_4^2}} \sim t(2)$$
$$②\frac{X_1}{2} \sim N(0,1)$$
$$\frac{X_2}{2} \sim N(0,1) \Rightarrow \frac{X_2^2}{4} \sim \chi^2(1)\text{且独立}$$
$$\frac{\frac{X_1}{2}}{\sqrt{\frac{X_2^2}{4}/1}} \sim t(1) \Rightarrow \frac{X_1}{|X_2|} \sim t(1)$$

设总体$X \sim N(\mu,\sigma^2)$, (X_1,X_2,X_3,X_4)为来自总体X的简单随机样本,求
$\dfrac{X_1-X_2}{|X_3-X_4|}$服从的分布.

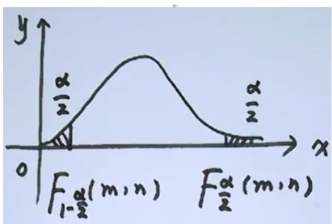
$$X_1-X_2 \sim N(0,2\sigma^2) \Rightarrow \frac{X_1-X_2}{\sqrt{2}\sigma} \sim N(0,1)$$
$$\frac{X_3-X_4}{\sqrt{2}\sigma} \sim N(0,1) \Rightarrow \frac{(X_3-X_4)^2}{2\sigma^2} \sim \chi^2(1)\text{且独立}$$
$$\frac{\frac{X_1-X_2}{\sqrt{2}\sigma}}{\sqrt{\frac{(X_3-X_4)^2}{2\sigma^2}/1}} \sim t(1) \Rightarrow \frac{X_1-X_2}{|X_3-X_4|} \sim t(1)$$

F分布

$$X \sim \chi^2(m), Y \sim \chi^2(n)\text{且}X, Y\text{独立,称}$$

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

$$\text{设}X \sim F(m,n)$$



性质

$$①X \sim t(n) \Rightarrow X^2 \sim F(1,n)$$

$$\text{证}: X \sim t(n) \Rightarrow$$

$$\exists U \sim N(0,1), V \sim \chi^2(n)\text{且}U, V\text{独立,使}$$

$$X = \frac{U}{\sqrt{V/n}}$$

$$\Rightarrow X^2 = \frac{U^2}{V/n}$$

$$\because U^2 \sim \chi^2(1), \text{且}U^2\text{与}V\text{独立}$$

$$\therefore X^2 = \frac{U^2/1}{V/n} \sim F(1,n)$$

$$②X \sim F(m,n) \Rightarrow \frac{1}{X} \sim F(n,m)$$

$$\text{总体}X \sim N(0,9), Y \sim N(0,9)\text{且独立}$$

$$(X_1, \cdots, X_9), (Y_1, \cdots, Y_9)$$

$$\frac{X_1^2 + \cdots + X_9^2}{Y_1^2 + \cdots + Y_9^2} \sim ?$$

$$\frac{X_i}{3} \sim N(0,1)(1 \le i \le 9)\text{且独立}$$

$$\Rightarrow \frac{1}{9}(X_1^2 + \cdots + X_9^2) \sim \chi^2(9)$$

$$\text{同理}\frac{1}{9}(Y_1^2 + \cdots + Y_9^2) \sim \chi^2(9)\text{且独立}$$

$$\frac{\frac{1}{9}(X_1^2 + \cdots + X_9^2)/9}{\frac{1}{9}(Y_1^2 + \cdots + Y_9^2)/9} \sim F(9,9)$$

正态总体下常见分布

$$X \sim N(\mu, \sigma^2) \Rightarrow (X_1, X_2, \cdots, X_n)$$

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$
$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$
$DX=\sigma^2-$总体方差
σ总体均方差
S^2样本方差$-$r.v.
S样本均方差

1.$\overline{X}\sim N(\mu,\dfrac{\sigma^2}{n})$

$\overline{X}=\dfrac{1}{n}X_1+\dfrac{1}{n}X_2+\cdots+\dfrac{1}{n}X_n$服从正态

$E\overline{X}=\dfrac{1}{n}\mu+\dfrac{1}{n}\mu+\cdots+\dfrac{1}{n}\mu=\mu$

$D\overline{X}=\dfrac{1}{n^2}\sigma^2+\cdots+\dfrac{1}{n^2}\sigma^2=\dfrac{\sigma^2}{n}$

$\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$

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

$\dfrac{X_1-\mu}{\sigma}\sim N(0,1),\cdots,\dfrac{X_n-\mu}{\sigma}\sim N(0,1)$且独立

3.$\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\sim\chi^2(n)$

$\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$

$\dfrac{n-1}{\sigma^2}\cdot\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\sim\chi^2(n-1)$

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

$E\dfrac{(n-1)S^2}{\sigma^2}=n-1\Rightarrow\dfrac{n-1}{\sigma^2}ES^2=n-1$

5.$ES^2=\sigma^2$

6.$\overline{X}$与S^2独立

设总体$X\sim N(0,4)$, $(X_1,X_2,\cdots,X_8)$为来自总体X的简单随机样本, $\overline{X}=\dfrac{1}{8}\sum_{i=1}^{8}X_i$,

$T=\sum_{i=1}^{8}(X_i-\overline{X})^2, a\overline{X}^2+bT\sim\chi^2(n)$, 求常数$a,b,n$.

$$1.\overline{X} \sim N(0,\frac{1}{2}) \Rightarrow \sqrt{2}\overline{X} \sim N(0,1) \Rightarrow 2\overline{X}^2 \sim \chi^2(1)$$

$$2.S^2 = \frac{1}{7}\sum_{i=1}^{8}(X_i - \overline{X})^2 \Rightarrow T = 7S^2$$

$$\frac{(8-1)S^2}{4} = \frac{T}{4} \sim \chi^2(7)$$

$3.\because \overline{X}$与S^2独立

$\therefore 2\overline{X}^2$与$\frac{T}{4}$独立

$$\therefore 2\overline{X}^2 + \frac{1}{4}T \sim \chi^2(8)$$

$$\therefore a = 2, b = \frac{1}{4}, n = 8$$

设总体X,Y独立同分布且都服从正态分布$N(0,9)$, $(X_1,\cdots,X_9)$与$(Y_1,\cdots,Y_9)$是分别来自总体X,Y的简单随机样本, 求统计量$U = \dfrac{X_1+X_2+\cdots+X_9}{\sqrt{Y_1^2+Y_2^2+\cdots+Y_9^2}}$所服从的分布.

$$X_1+\cdots+X_9 \sim N(0,81) \Rightarrow \frac{1}{9}(X_1+\cdots+X_9) \sim N(0,1)$$

$\frac{Y_i}{3} \sim N(0,1)(1 \le i \le 9)$且独立

$$\Rightarrow \frac{1}{9}(Y_1^2+\cdots+Y_9^2) \sim \chi^2(9)$$

$$\frac{\frac{1}{9}(X_1+\cdots+X_9)}{\sqrt{\frac{1}{9}(Y_1^2+\cdots+Y_9^2)/9}} \sim t(9)$$

设$X_i \sim N(0,4)(1 \le i \le 6)$且$X_1,X_2,\cdots,X_6$相互独立, 令$S_0^2 = \dfrac{1}{4}\sum\limits_{i=3}^{6}X_i^2$, 求$\dfrac{X_1+X_2}{\sqrt{2}S_0}$所服从的分布.

$$X \sim N(0,4) \Rightarrow (X_1,\cdots,X_6)$$

$$X_1+X_2 \sim N(0,8)$$

$$\Rightarrow \frac{X_1+X_2}{2\sqrt{2}} \sim N(0,1)$$

$\frac{X_i}{2} \sim N(0,1)(3 \le i \le 6)$且独立

$\frac{1}{4}(X_3^2+\cdots+X_6^2) \sim \chi^2(4)$, 即$S_0^2 \sim \chi^2(4)$

而$\frac{X_1+X_2}{\sqrt{2}}$与S_0^2独立

$$\frac{\frac{X_1+X_2}{2\sqrt{2}}}{\sqrt{S_0^2/4}} \sim t(4) \Rightarrow \frac{X_1+X_2}{\sqrt{2}S_0} \sim t(4)$$