第六章 样本及抽样分布

§1 随机样本

§2 抽样分布

第六章 样本及抽样分布

§1 随机样本

- 总体
- 个体
- 样本

一、总体和个体

1）总体：研究对象的某项数量指标的值的全体。

2）个体：总体中的每个元素为个体。

例如：某工厂生产的灯泡的寿命是一个总体，每一个灯泡的寿命是一个个体；某学校男生的身高的全体一个总体，每个男生的身高是一个个体。

二、样本

定义：设 X 是具有分布函数 F 的随机变量，若 $X_1,\cdots X_n$ 是具有同一分布函数 F 的相互独立的随机变量，则称 $X_1,\cdots X_n$ 为从总体 X 中得到的容量为 n 的简单随机样本，简称为样本，其观察值 $x_1,\cdots x_n$ 称为样本值。

由定义知：若 $X_1,\cdots,X_n$ 为 X 的一个样本，则 $(X_1,\cdots,X_n)$ 的联合分布函数为：

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

若设X的概率密度为 $f(x)$，则 $(X_1,\cdots,X_n)$ 的联合概率密度为：

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

若设X的分布律为 $P\{X=x\}=p(x)$ ，则 $(X_1,\cdots,X_n)$ 的联合分布律为：

$$P\{X_1=x_1,\cdots,X_n=x_n\}=\prod_{i=1}^{n}p(x_i)$$

例1 若 $X_1,\cdots,X_n$ 是正态总体 $X\sim N(1,4)$ 的样本，则

$EX_1X_n=$ $\underline{\ 1\ }$， $D(X_1-2X_2)=$ $\underline{\ 20\ }$.

例2 若$X_1,\cdots,X_n$是总体$X\sim B(1,p)$的样本，求$(X_1,\cdots,X_n)$的联合分布律.

解： 总体 X的分布律为

$$p(x)=P\{X=x\}=p^x(1-p)^{1-x},x=0,1.$$

所以（$X_1,\cdots,X_n$）的联合分布律为

$$P\{X_1=x_1,\cdots,X_n=x_n\}=\prod_{i=1}^{n}p(x_i)$$

$$=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i}$$

$$x_i=0,1,\ i=1,\cdots,n.$$

例3 若$X_1,\cdots,X_n$是参数为λ的泊松分布总体X的样本，

求$(X_1,\cdots,X_n)$的联合分布律

解： 总体 X 的分布律为

$$p(x)=P\{X=x\}=\frac{\lambda^x}{x!}e^{-x},x=0,1,\cdots$$

所以（$X_1,\cdots,X_n$）的联合分布律为

$$P\{X_1=x_1,\cdots,X_n=x_n\}=\prod_{i=1}^{n}p(x_i)$$

$$=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-x_i}=\frac{\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}x_i!}e^{-\sum_{i=1}^{n}x_i},\quad x_i=0,1,\cdots,\quad i=1,\cdots,n.$$

例4 若$X_1,\cdots,X_n$是总体$X\sim N(\mu,\sigma^2)$的样本，求$(X_1,\cdots,X_n)$的联合概率密度

解: 总体 X 的概率密度为

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<\infty.$$

所以（$X_1,\cdots,X_n$）的联合概率密度为

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

$$=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}=(2\pi)^{-\frac{n}{2}}e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}}$$

$$-\infty<x_i<\infty,i=1,\cdots,n.$$

§2 抽样分布

统计量

χ^2-分布

$t-$分布

$F-$分布

正态总体的样本均值与样本方差的分布

一、统计量及其分布

1） 定义：设 $X_1,\cdots X_n$ 为来自总体X的一个样本，g是 $X_1,\cdots X_n$ 的函数，若g是连续函数，且g中不含任何未知参数， 则称 $g(X_1,\cdots X_n)$ 是统计量。

设 $(x_1,\cdots,x_n)$ 是相应于样本 $(X_1,\cdots X_n)$ 的样本值。

则称 $g(x_1,\cdots x_n)$ 是 $g(X_1,\cdots X_n)$ 的观察值。

注：统计量是随机变量。

例1 设 $X_1,\cdots X_n$ 为来自总体 $X\sim N(\mu,\sigma^2)$ 的一个样本，其中 μ 未知，σ^2 已知，问下列随机变量中那些是统计量

$$\min(X_1,X_2,\cdots,X_n);\quad \frac{X_1+X_n}{2};\quad \frac{X_1+\cdots+X_n}{n}-\mu;$$

$$\frac{(X_1+X_n)^2}{\sigma^2};\quad \frac{(X_1+\cdots+X_n)-n\mu}{\sqrt{n}\sigma}.$$

2） 常用的统计量

样本均值 $\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i,$

样本方差 $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\dfrac{1}{n-1}[\sum_{i=1}^{n}X_i^2-n\overline{X}^2]$

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

证明:

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

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

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

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

样本标准差 $S=\sqrt{S^2}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$

样本k阶原点矩 $A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^{\,k} \quad k=1,2,\cdots$

样本k阶中心矩 $B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k \quad k=1,2,\cdots$

它们的观察值分别为：

$$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^{\,2}-n\overline{x}^2\right]$$

$$s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$a_k=\frac{1}{n}\sum_{i=1}^{n}x_i^{\ k},\quad k=1,2\cdots$$

$$b_k=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^k,\quad k=1,2\cdots$$

分别称为样本均值、样本方差、样本标准差、样本k阶原点矩、样本k阶中心矩的观察值。

统计量是样本的函数，它是一个随机变量，统计量的分布称为抽样分布。

3）结论：设 $X_1,\cdots X_n$ 为来自总体X 的一个样本，

$$EX=\mu,\ DX=\sigma^2,$$

则

$$E\overline{X}=\mu,\ D\overline{X}=\frac{\sigma^2}{n},\ ES^2=\sigma^2.$$

请记熟此结论！

证明：

$$E\overline{X}=E\left(\frac{\sum_{i=1}^{n}X_i}{n}\right)=\frac{\sum_{i=1}^{n}EX_i}{n}=\mu$$

$$D\overline{X}=D\left(\frac{\sum_{i=1}^{n}X_i}{n}\right)=\frac{\sum_{i=1}^{n}DX_i}{n^2}=\frac{\sigma^2}{n}$$

$$ES^2 = E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2] = \frac{1}{n-1}E[\sum_{i=1}^{n} X_i^{\ 2} - n\overline{X}^{\,2}]$$

$$= \frac{1}{n-1}[\sum_{i=1}^{n} EX_i^{\ 2} - nE\overline{X}^{\,2}]$$

$$= \frac{1}{n-1}[\sum_{i=1}^{n}(DX_i + (EX_i)^2) - n(D\overline{X} + (E\overline{X})^2)]$$

$$= \frac{1}{n-1}[\sum_{i=1}^{n}(\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2)]$$

$$= \frac{1}{n-1}(n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2) = \sigma^2$$

二、常用统计量的分布

1) χ^2-分布

设$(X_1,\cdots X_n)$为来自于正态总体$N(0,1)$的样本，

则称统计量：$\chi^2=X_1^2+\cdots+X_n^2$

所服从的分布为自由度是n的χ^2分布。

记为 $\chi^2\sim\chi^2(n)$

χ^2分布的性质：

1^0若 $X\sim\chi^2(n_1),Y\sim\chi^2(n_2)$,且$X$，$Y$独立，则有

$$X+Y\sim\chi^2(n_1+n_2)$$

证明: 设 $Z = X + Y$,

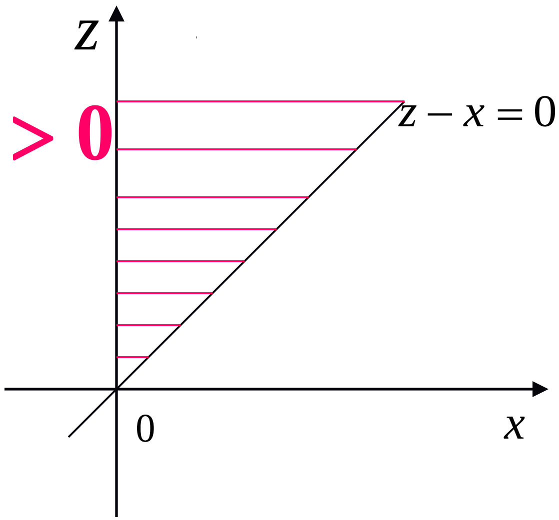
由于

$$f_X(x) = \begin{cases} \dfrac{1}{2^{\frac{m}{2}}\Gamma\left(\dfrac{m}{2}\right)} x^{\frac{m}{2}-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & x \le 0 \end{cases}$$

$$f_Y(y) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} & y > 0 \\ 0 & y \le 0 \end{cases}$$

设随机变量$Z=X+Y$的密度函数为$f_Z(z)$，则有

$$f_Z(z)=\int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)dx \qquad x>0,\ z-x>0$$



(1) 当 $z\le 0$，$f_z(z)=0$.

(2) 当 $z>0$，$f_z(z)=$

$$=\int_0^z \frac{1}{2^{\frac{m}{2}}\Gamma\left(\frac{m}{2}\right)} x^{\frac{m}{2}-1} e^{-\frac{x}{2}} \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} (z-x)^{\frac{n}{2}-1} e^{-\frac{z-x}{2}} dx$$

$$f_Z(z) = \frac{e^{-\frac{z}{2}}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\int_0^z x^{\frac{m}{2}-1}(z-x)^{\frac{n}{2}-1}dx$$

$$= \frac{e^{-\frac{z}{2}}z^{\frac{n}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\int_0^z x^{\frac{m}{2}-1}\left(1-\frac{x}{z}\right)^{\frac{n}{2}-1}dx$$

作积分变换$t=\dfrac{x}{z}$，$dt=\dfrac{dx}{z}$

当$x=0$时，$t=0$；　当$x=z$时，$t=1$.

$$f_Z(z)=\frac{e^{-\frac{z}{2}}z^{\frac{n}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\int_0^1 (tz)^{\frac{m}{2}-1}(1-t)^{\frac{n}{2}-1}z\,dt$$

$$=\frac{e^{-\frac{z}{2}}z^{\frac{m+n}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\int_0^1 t^{\frac{m}{2}-1}(1-t)^{\frac{n}{2}-1}dt$$

由数学中$B-$函数的定义：

$$B(s,\ t)=\int_0^1 x^{s-1}(1-x)^{t-1}dx\quad(s>0,\ t>0)$$

以及$B-$函数与$\Gamma-$函数之间的关系：$B(s,\ t)=\dfrac{\Gamma(s)\Gamma(t)}{\Gamma(s+t)}$

可知， $f_Z(z)=\dfrac{e^{-\frac{z}{2}}z^{\frac{m+n}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\cdot B\left(\frac{m}{2},\ \frac{n}{2}\right)$

$$=\frac{e^{-\frac{z}{2}}z^{\frac{m+n}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\cdot\frac{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{m+n}{2}\right)}=\frac{e^{-\frac{z}{2}}z^{\frac{m+n}{2}-1}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)}$$

综上所述，我们有

$$f_Z(z)=\begin{cases}\dfrac{1}{2^{\frac{m+n}{2}}\Gamma\left(\dfrac{m+n}{2}\right)}e^{-\frac{z}{2}}z^{\frac{m+n}{2}-1} & z>0\\ 0 & z\le 0\end{cases}$$

由此，我们得

如果随机变量 X 与 Y 相互独立，且

$$X\sim\chi^2(m),\ Y\sim\chi^2(n),$$

$Z=X+Y,$

则 $$Z\sim\chi^2(m+n)$$

2^0 若 $\chi^2 \sim \chi^2(n)$, 则 $E\chi^2 = n,\quad D\chi^2 = 2n.$

证: $\chi^2 = X_1^2 + \cdots + X_n^2$

$$X_i \sim N(0,1),\quad EX_i = 0,\quad DX_i = 1,$$

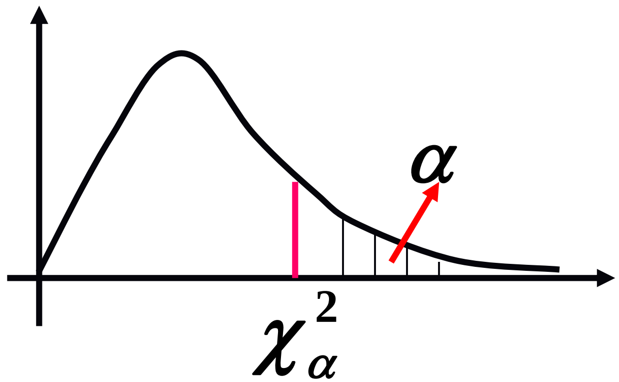
$$EX_i^2 = DX_i + (EX_i)^2 = 1,$$

$$DX_i^2 = EX_i^4 - (EX_i^2)^2 = 3 - 1 = 2,\quad i = 1,2,\cdots n$$

所以 $E\chi^2 = E(\sum_{i=1}^{n} X_i^2) = \sum_{i=1}^{n} EX_i^2 = n$

$$D\chi^2 = D(\sum_{i=1}^{n} X_i^2) = \sum_{i=1}^{n} DX_i^2 = 2n$$

返回主目录



对于给定的$\alpha(0<\alpha<1)$，称满足条件：

$$P\{\chi^2>\chi_\alpha^2(n)\}=\alpha$$

的点$\chi_\alpha^2(n)$为$\chi^2(n)$分布的上α分位点。

当n充分大时，$\chi_\alpha^2(n)\approx\frac{1}{2}(z_\alpha+\sqrt{2n-1})^2$

z_α是标准正态分布的上α分位点。

例2

设$(X_1,\cdots X_n)$为来自于正态总体$N(\mu,\sigma^2)$的样本，

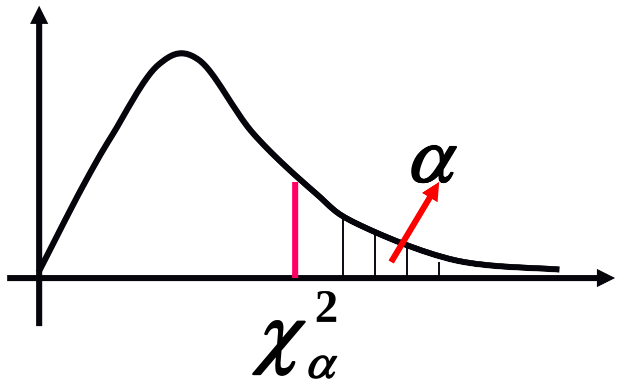
则 $\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}\sim$ ___$\chi^2(n)$___.

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

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

例3 $\chi^2_{0.05}(8)=$ ___15.507___, $\chi^2_{0.95}(8)=$ ___2.733___.

$\chi^2_{0.05}(50)=$ ___67.22___, $\chi^2_{0.05}(100)=$ ___124.06___.



对于给定的$\alpha(0<\alpha<1)$，称满足条件：

$$P\{\chi^2>\chi_\alpha^2(n)\}=\alpha$$

的点$\chi_\alpha^2(n)$为$\chi^2(n)$分布的上α分位点。

当n充分大时，$\chi_\alpha^2(n)\approx\frac{1}{2}(z_\alpha+\sqrt{2n-1})^2$

z_α是标准正态分布的上α分位点。

例4

若 $X \sim \chi^2(6)$，且 λ_1 使 $P\{X > \lambda_1\} = 0.05$，则

$\lambda_1 = \underline{\chi^2_{0.05}(6) = 12.592}$.

若 $X \sim \chi^2(9)$，且 λ_2 使 $P\{X < \lambda_2\} = 0.05$，则

$\lambda_2 = \underline{\chi^2_{0.95}(9) = 3.325}$.

2) $t-$分布

$X \sim N(0,1), Y \sim \chi^2(n)$, X, Y独立，则称随机变量

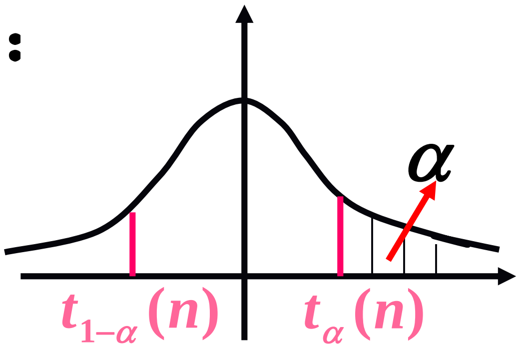
$$t = \frac{X}{\sqrt{Y/n}}$$

所服从的分布为自由度是n的$t-$分布，记作$t \sim t(n)$.

对于给定的$\alpha(0<\alpha<1)$，称满足条件：

$$P\{t > t_\alpha(n)\} = \alpha$$

的点$t_\alpha(n)$为t分布的上α分位点。



由概率密度的对称性知：$t_{1-\alpha}(n) = -t_\alpha(n)$

当$n>45$时，$t_\alpha(n) \approx z_\alpha$

例5 $t_{0.05}(8) = \underline{1.8595}$, $t_{0.95}(8) = \underline{-1.8595}$.

$t_{0.05}(50) = \underline{1.645}$, $t_{0.95}(50) = \underline{-1.645}$.

例6

若 $X \sim t(10)$,且 λ_1 使 $P\{X > \lambda_1\} = 0.05$,则

$\lambda_1 = \underline{t_{0.05}(10) = 1.8125}$.

若 $X \sim t(9)$,且 λ_2 使 $P\{X < \lambda_2\} = 0.05$,则

$\lambda_2 = \underline{t_{0.95}(9) = -1.8331}$.

3) F－分布

若 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，$X,Y$ 独立，则称随机变量

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

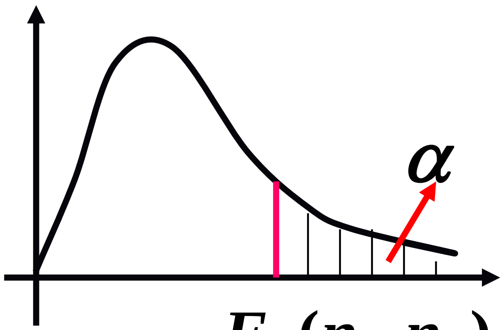
所服从的分布为自由度是 n_1, n_2 的 F－分布，记作 $F \sim F(n_1, n_2)$.

定理：若 $F \sim F(n_1, n_2)$，则 $1/F \sim F(n_2, n_1)$.

对于给定的 $\alpha(0<\alpha<1)$，称满足条件：

$$P\{F > F_\alpha(n_1, n_2)\} = \alpha$$

的点 $F_\alpha(n_1, n_2)$ 为 F 分布的上 α 分位点。



结论：$F_{1-\alpha}(n_1, n_2) = 1/F_\alpha(n_2, n_1)$

返回主目录

$$P\{F > F_\alpha(n_1,n_2)\} = \alpha$$

证明：若$F \sim F(n_1,n_2)$

$$1-\alpha = P\{F > F_{1-\alpha}(n_1,n_2)\} = P\{\frac{1}{F} < \frac{1}{F_{1-\alpha}(n_1,n_2)}\}$$

$$= 1 - P\{\frac{1}{F} \geq \frac{1}{F_{1-\alpha}(n_1,n_2)}\}$$

所以 $P\{\frac{1}{F} > \frac{1}{F_{1-\alpha}(n_1,n_2)}\} = \alpha$

又因为 $1/F \sim F(n_2,n_1)$, 所以 $F_\alpha(n_2,n_1) = \frac{1}{F_{1-\alpha}(n_1,n_2)}$

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

例7 $F_{0.05}(8,6)=\underline{4.15}$, $F_{0.95}(8,6)=\underline{\dfrac{1}{F_{0.05}(6,8)}=0.28}$.

$F_{0.1}(4,7)=\underline{2.96}$, $F_{0.9}(4,7)=\underline{\dfrac{1}{F_{0.1}(7,4)}=0.25}$.

例8

若 $X\sim F(10,6)$, 且 λ_1 使 $P\{X>\lambda_1\}=0.05$, 则

$\lambda_1=\underline{F_{0.05}(10,6)=4.06}$.

若 $X\sim F(9,8)$, 且 λ_2 使 $P\{X<\lambda_2\}=0.05$, 则

$\lambda_2=\underline{F_{0.95}(9,8)=\dfrac{1}{F_{0.05}(8,9)}=0.31}$.

例9 已知 $X \sim t(n)$, 试证 $X^2 \sim F(1,n)$.

解: 由于 $X \sim t(n)$, 所以 $X = \dfrac{Y}{\sqrt{Z/n}}$,

其中 $Y \sim N(0,1)$, $Z \sim \chi^2(n)$, 且 Y, Z 独立.

则 $X^2 = \dfrac{Y^2/1}{Z/n}$,

由 F 分布的定义知 $X^2 \sim F(1,n)$.

4）正态总体的样本均值与样本方差的分布：

定理1 设$X_1,\cdots,X_n$是总体$N(\mu,\sigma^2)$的样本，$\bar{X},S^2$分别是样本均值与样本方差，则有：

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

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

(3) $\bar{X}$与S^2独立。

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)$P\{S^2 > 1.85\}$;

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

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

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

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

例10（续）

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

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

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

（3） 由于 ${(9-1)S^2}/{9}\sim\chi^2(8)$ ，故

$$P\{S^2>1.85\}=P\{{8S^2}/{9}>1.644\}\approx 0.99$$

例10（续）

（4） $P\{\max(X_1,X_2,X_3)>4\}$

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

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

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

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

$X_1\sim N(2,9)$

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

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

$$=0.58$$

例10（续）

（5） $P\{\min(X_1, X_2, X_3) < 0\}$

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

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

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

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

$X_1 \sim N(2,9)$

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

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

$$= 0.58$$

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

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

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

$$\frac{(n-1)S^2}{\sigma^2}\sim \chi^2(n-1),$$

且它们独立。

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

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

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。

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

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

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

返回主目录

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

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

返回主目录

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 $P\{|\bar{X}-\bar{Y}|>0.1\}$.

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

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

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

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

$$= 0.8886$$

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。

则：（1）
$$\frac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2/\sigma_1^2}{\sum_{j=1}^{n_2}(Y_j-\mu_2)^2/\sigma_2^2}\cdot\frac{n_2}{n_1} \sim F(n_1, n_2)$$

（2）
$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\sum_{i=1}^{n_1}(X_i-\overline{X})^2/\sigma_1^2}{\sum_{j=1}^{n_2}(Y_j-\overline{Y})^2/\sigma_2^2}\cdot\frac{n_2-1}{n_1-1}$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$$\sim F(n_1-1, n_2-1)$$

例12

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

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

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

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

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

例12（续）

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

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

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

例12（续）

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

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

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

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。