Chpt.6 Sampling and Distribution

第六章 样本及抽样分布

上节回顾

- 概率论 vs 数理统计

概率论：理论，已知服从某种分布，研究性质与规律

数理统计：实践，收集数据，分析数据，做统计推断

抽样，简单随机抽样

- 总体X

样本$X_1,X_2,...,X_n$,独立同分布（与X同分布）

统计量$g(X_1,X_2,...,X_n)$

- 样本均值 $\overline{X}=\dfrac{1}{n}\sum_{k=1}^{n}X_k$
- 样本方差 $S^2=\dfrac{1}{n-1}\sum_{k=1}^{n}(X_i-\overline{X})^2$

$\overline{X}\xrightarrow{P}E(X)$ $\quad E(\overline{X})=E(X)$ $\quad D(\overline{X})=\frac{1}{n}D(X)$

$S^2\xrightarrow{P}D(X)$ $\quad E(S^2)=D(X)$

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

$$A_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

$$E(S^2)=D(X)$$

$$E(A_2)=\frac{n-1}{n}D(X)$$

统计学中，用样本统计量来估计总体参数，例如用样本均值$\overline{X}$来估计总体的期望E(X)。在估计时，一般要求为无偏估计，即估计量的数学期望等于被估计参数的真实值。 样本方差时对总体方差的无偏估计。

设$X_1, X_2, \cdots, X_n$是总体X的简单随机样本，S^2是样本方差。

证明：$E(S^2)=D(X)$

证明：设 $E(X)=\mu$，$D(X)=\sigma^2$.

$X_1, \cdots, X_n$ 独立同分布，与X的分布相同.

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

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

$D(X_i)=D(X)$

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

[注]：$S^2=\frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-(\overline{X}-\mu)^2\right]$.

由大数定律，$\overline{X}\xrightarrow{P}\mu$，$(\overline{X}-\mu)^2\xrightarrow{P}0$

$\{(X_i-\mu)^2\}$独立同分布，$E(X_i-\mu)^2=D(X_i)=\sigma^2$，由大数定律，$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\xrightarrow{P}\sigma^2$

$\therefore S^2\xrightarrow{P}\sigma^2$，$n\to\infty$.

经验分布函数　我们还可以作出与总体分布函数 $F(x)$ 相应的统计量——经验分布函数. 它的作法如下：设 $X_1, X_2, \cdots, X_n$ 是总体 F 的一个样本，用 $S(x)$，$-\infty<x<\infty$ 表示 $X_1, X_2, \cdots, X_n$ 中不大于 x 的随机变量的个数. 定义经验分布函数 $F_n(x)$ 为

$$F_n(x)=\frac{1}{n}S(x),\quad -\infty<x<\infty.$$

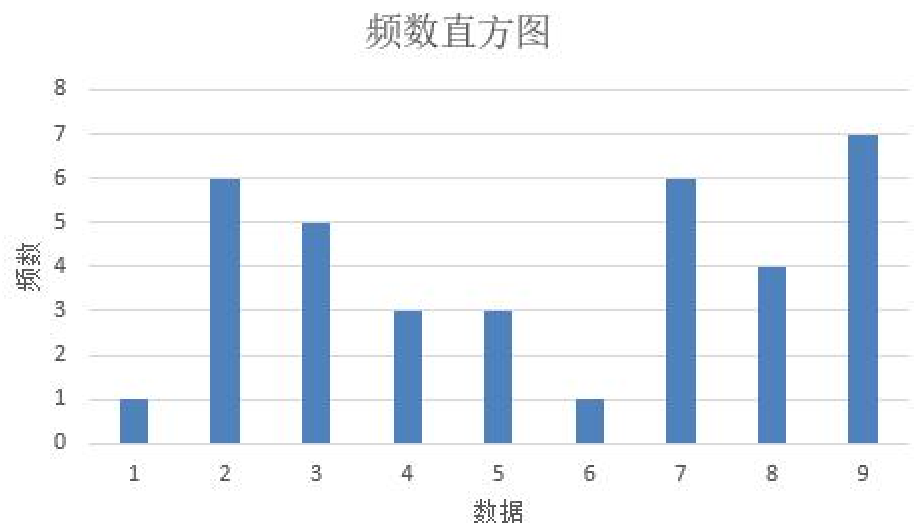
可以证明：

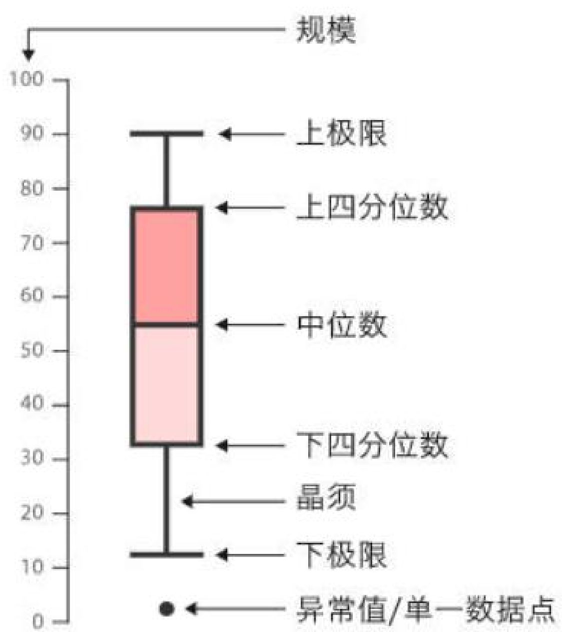
对于任一实数 x，当 $n\to\infty$ 时 $F_n(x)$ 以概率 1 一致收敛于分布函数 $F(x)$，即

$$P\{\lim_{n\to\infty}\sup_{-\infty<x<\infty}|F_n(x)-F(x)|=0\}=1.$$

因此，对于任一实数 x 当 n 充分大时，经验分布函数的任一个观察值 $F_n(x)$ 与总体分布函数 $F(x)$ 只有微小的差别，从而在实际上可当作 $F(x)$ 来使用①.

直方图，箱形图





见教材130-135页，请同学们自学

6.3 抽样分布

样本是随机变量

统计量是样本的函数，从而统计量也是随机变量

统计量的分布称为抽样分布

为什么要研究抽样分布：

- 一般而言，总体分布已知，抽样分布也是知道的，但是确切得到是困难的；
- 从另外一个角度，我们希望由统计量的分布（特别是在观测值得到后），估计、推断出总体的一些特征。

一.样本均值分布

假设总体X分布的均值与方差都是已知的，那么可以对来自总体的多个样本$X_1,X_2...,X_n$的均值$\overline{X}$做出估计。

$$E(\overline{X})=E(X)$$

$$D(\overline{X})=\frac{1}{n}D(X)$$

假设$X_1,X_2...,X_n$是来自总体X的独立样本, 样本均值为 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 由中心极限定理可知，当n充分大时，

$$\overline{X}\overset{\text{近似地}}{\sim}N(\mu,\frac{\sigma^2}{n}) \qquad \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\overset{\text{近似地}}{\sim}N(0,\ 1)$$

假设 $X_1,\cdots,X_n$ 是来自正态总体$X\sim N(\mu,\sigma^2)$ 的独立样本,

则样本均值 $\overline{X}$ 服从正态分布 $\overline{X}\sim N(\mu,\frac{\sigma^2}{n})$

标准量服从标准正态分布 $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

例题1：从正态分布总体$N(3.4, 6^2)$中抽取容量为n的样本，如果要求其样本均值位于区间（1.4，5.4）内的概率不小于0.95，问样本容量n至少应取多大？

z	1.28	1.645	1.96	2.33
$\Phi(z)$	0.900	0.950	0.975	0.990

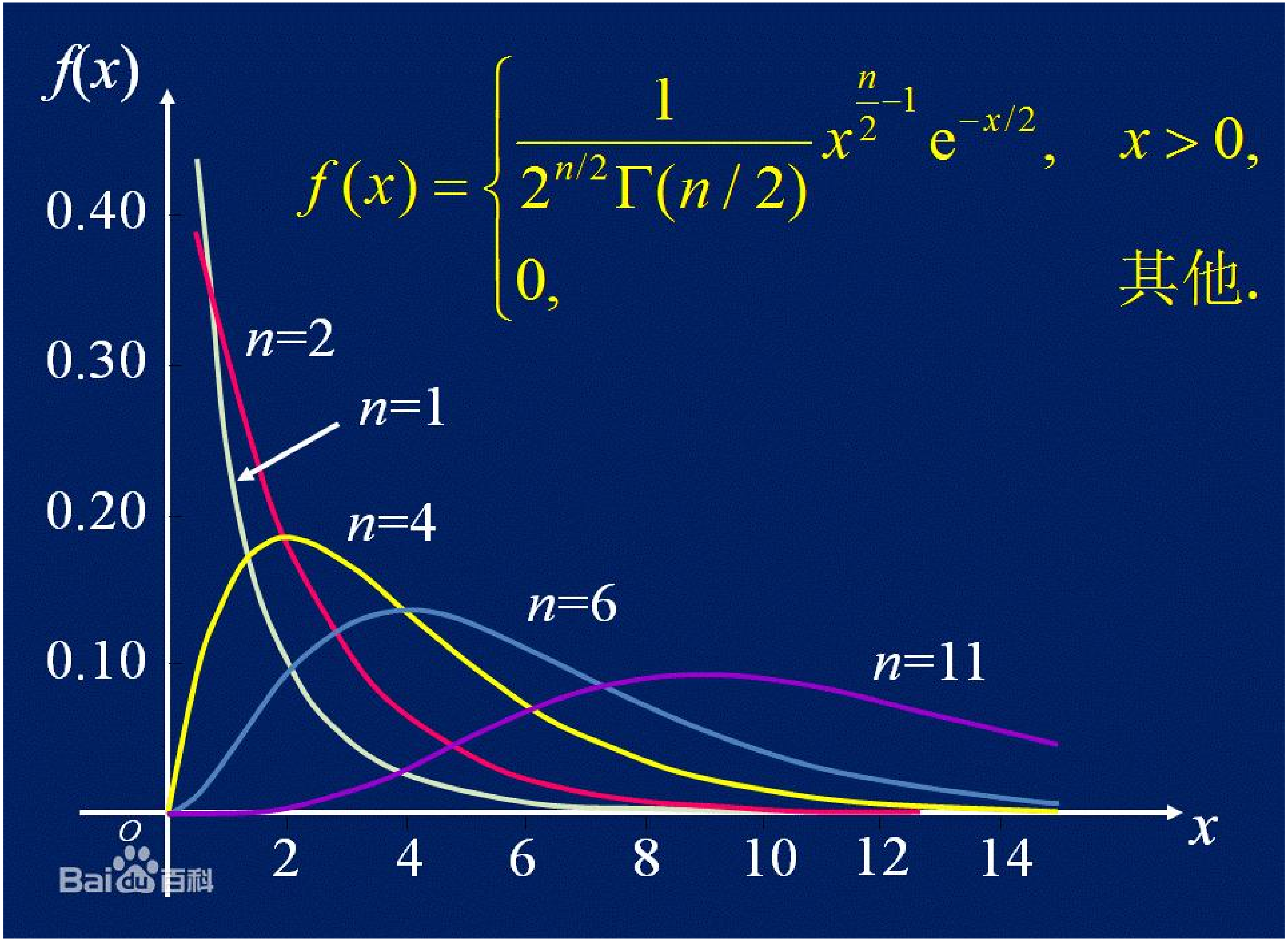
二. χ^2 分布

[定理6.3] 设随机变量 $X_1,\cdots,X_n$ 是来自标准正态总体

$X \sim N(0,1)$ 的独立样本。则随机变量 $\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$

服从自由度为 n 的 χ^2分布，记为 $\chi^2(n)$，其概率密度：

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



$f(x)$
$$f(x)=\begin{cases}\dfrac{1}{2^{n/2}\Gamma(n/2)}x^{\frac{n}{2}-1}e^{-x/2}, & x>0,\\ 0, & 其他.\end{cases}$$
0.40
0.30
0.20
0.10
$n=2$
$n=1$
$n=4$
$n=6$
$n=11$
O
2
4
6
8
10
12
14
x

Remark 1： χ^2分布具有可加性，也就是说，

$\chi_1^2 \sim \chi^2(n_1)$, $\chi_2^2 \sim \chi^2(n_2)$ 且它们相互独立，

则 $\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$

Remark 2： $\chi^2 \sim \chi^2(n)$

则 $E(\chi^2) = n$ ， $D(\chi^2) = 2n$

证明$D(\chi^2(n))$ =2n

- 设X~N(0,1), 若证明$D(X^2)=2$，则可说明 $D(\chi^2(n))$ =2n

用平方关系来算, $D(X^2)=E(X^4)-(E(X^2))^2$
先算 $E(X^4)$, 令$f(x)$ 是N (0,1)的密度函数

思考：利用上述方法，求 $E(X^k)$，$D(X^k)$

$$E(X^4)=\int_{-\infty}^{+\infty}x^4 f(x)dx=\int_{-\infty}^{+\infty}x^3\cdot xf(x)dx$$

$$f'(x)=f(x)\left(-\frac{x^2}{2}\right)'=-xf(x)$$

$$E(X^4)=\int_{-\infty}^{+\infty}-x^3\cdot f'(x)dx=-x^3\cdot f(x)|_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty}-3x^2\cdot f(x)dx$$

$$=0-\int_{-\infty}^{+\infty}-3x\cdot xf(x)dx=-\int_{-\infty}^{+\infty}3x\cdot f'(x)dx$$

$$=-3x\cdot f(x)|_{-\infty}^{+\infty}+3\int_{-\infty}^{+\infty}f(x)dx=3$$

$$E(X^2)=D(X)+E^2(X)=1+0=1 \text{ 所以 } (E(X^2))^2=1$$

$$\text{则}D(X^2)=E(X^4)-(E(X^2))^2=3-1=2$$

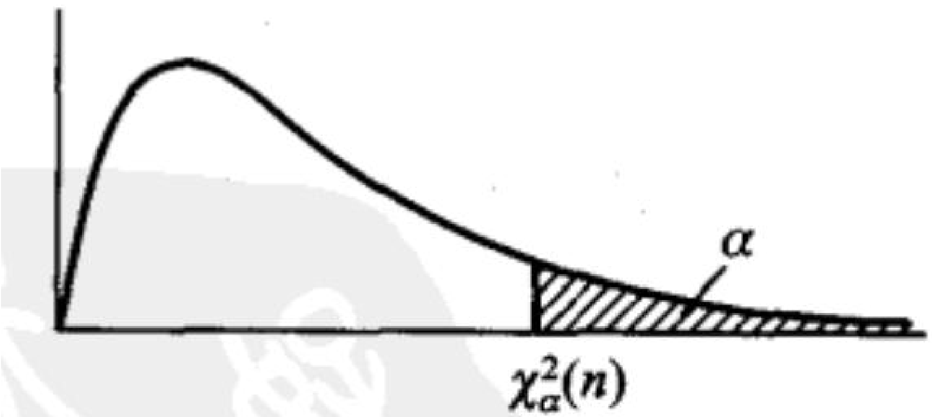
[定义] 随机变量X, 对一个正数α (0< α <1) , 满足

$$P\{X > x_\alpha\} = \alpha$$ 的值 X_α 称为X分布的上α分位点.

对于实数 $\alpha(0<\alpha<1)$，称满足下式

$$P\{\chi^2>\chi^2_\alpha(n)\}=\int_{\chi^2_\alpha(n)}^{\infty}f(x)dx=\alpha$$

的数 $\chi^2_\alpha(n)$ 为χ^2分布的上α分位点。



- 对于n≤40，χ^2分布的上α分位点的值可参考教材的附表5
- 对于n>40，χ^2分布的上α分位点可以由下面的近似关系求解

$$\chi^2_\alpha(n)\approx\frac{1}{2}(z_\alpha+\sqrt{2n-1})^2$$

其中 z_α 是标准正态分布的上 α 分位点

假设正态分布X~N(μ，σ²)，n个样本 $X_1, \cdots, X_n$ ，则

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

(注意：χ^2分布要求X~N(0,1))

- μ已知，可估计σ；或σ已知，可估计μ。
- 当均值μ与方差σ都未知时，该如何估计？

考虑用 $(n-1)S^2 = \sum_{i=1}^{n}(X_i-\bar{X})^2$ 来代替 $\sum_{i=1}^{n}(X_i-\mu)^2$

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

[定理6.4] S²是来自正态总体 $X \sim N(\mu,\sigma^2)$ 的n个样本的方差，那么：

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

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

例题2：设$X_1,X_2,...,X_{10}$是来自标准正态分布的一组简单随机样本，

$$Y=\frac{1}{2}\sum_{i=1}^{10}X_i^2+\sum_{i=1}^{5}X_{2i-1}X_{2i}$$

则EY=____，Y服从_____分布，参数是_______

三. T分布（回顾中心极限定理 VS. T分布）

当总体均值与方差已知 $E(X)=\mu, D(X)=\sigma^2$ ，$\{X_i\}$ 是来自总体的独立样本，我们知道 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ 近似为标准正态分布 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \rightarrow N(0,1)$

如果X是正态分布，那就可以确切得到 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,\ 1)$

- 已知σ，可以估计μ；反过来, 已知μ，可以估计σ
- 那如果μ和σ 都未知，该如何估计？

考虑用样本标准差S来代替σ

三. T分布

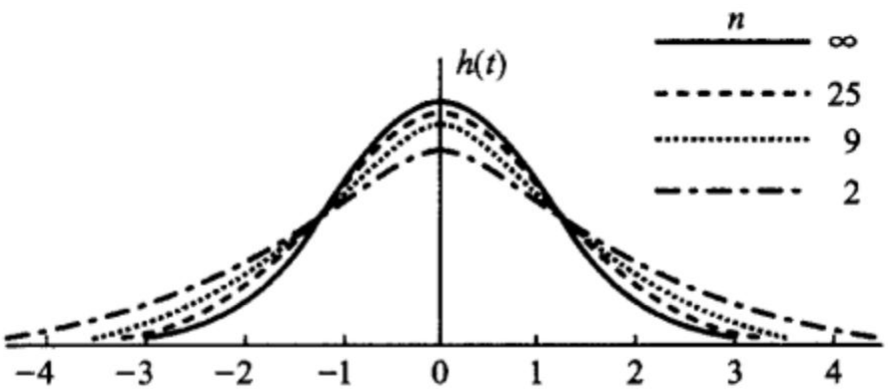
[定义] 设 $X \sim N(0,1)$, $Y \sim \chi^2(n)$, 且 X,Y 相互独立, 则称随机变量

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

服从自由度为 n 的 t 分布. 记为 $t \sim t(n)$.

t 分布又称**学生氏**(Student)**分布**. $t(n)$分布的概率密度函数为

$$h(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\ \Gamma(n/2)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}, -\infty < t < \infty$$



Remark：

- t 分布的密度曲线关于 $t = 0$ 对称；
- n充分大时，近似于标准正态分布。

$$\lim_{n\to\infty} h(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

三. T分布

[定理6.1] 若总体服从正态分布 $X \sim N(\mu,\sigma^2)$ ，$\{X_i\}, S$ 分别是来自总体的样本与样本标准差, $S^2 = \dfrac{1}{n-1}\sum_{k=1}^{n}(X_i - \overline{X})^2$ 那么随机变量 $t = \dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ 服从自由度为n-1的t分布，记为 $t \sim t(n-1)$ ，其概率密度为：

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

Remark1: 其中伽玛函数 $\Gamma(\alpha)=\int_0^{\infty} u^{\alpha-1}e^{-u}du \quad (\alpha>0)$，有如下性质：

(1) $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$

(2) $\Gamma(n)=(n-1)!$

(3) $\Gamma(\frac{1}{2})=\sqrt{\pi}$

Remark2: 由定理可以知道总体为正态分布 $X\sim N(\mu,\sigma^2)$，随机变量 $t=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ 服从自由度为n-1的 t 分布。

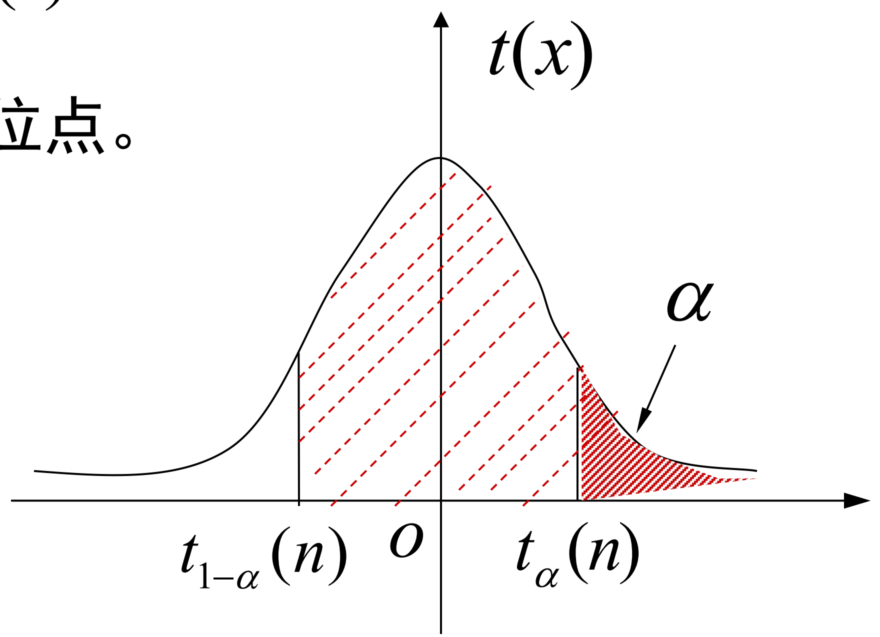
$t=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ 中如果μ和σ未知，由对 t 的分析可以估计出μ，从而估计σ。

T分布的分位点

对于给定的α (0< α <1) , 满足条件

$$P\{t > t_\alpha(n)\} = \int_{t_\alpha(n)}^{\infty} f_t(t)dt = \alpha$$

的 $t_\alpha(n)$ 值称为t分布的上α分位点。



由 $f_t(t)$ 的对称性知：

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

t 分布的上 α 分位点可自附表 4 查得. 在 $n>45$ 时，对于常用的 α 的值，就用正态近似　　$t_\alpha(n) \approx z_\alpha$.

例题3：设X和Y相互独立，都服从正态分布N(0,9)，而 $X_1, X_2, \ldots, X_9$ 和 $Y_1, Y_2, \ldots, Y_9$ 分别是来自总体X和Y的简单随机样本，则统计量

$$U = \frac{X_1 + X_2 + \ldots + X_9}{\sqrt{Y_1^2 + Y_2^2 + \ldots + Y_9^2}}$$

服从什么分布，参数为多少？

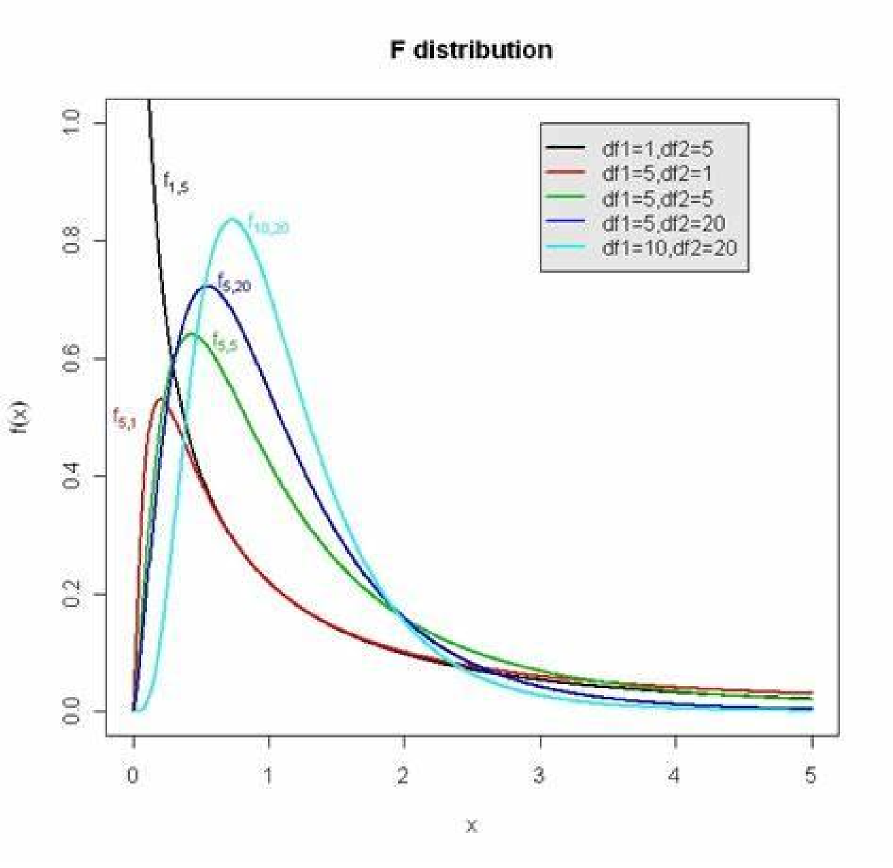
四. F分布

[定义] 若随机变量 $U \sim \chi^2(n_1), V \sim \chi^2(n_2)$, U与V独立, 则随机变量 $F = \dfrac{U/n_1}{V/n_2}$ 服从自由度为(n_1, n_2) 的 F 分布，记为 $F \sim F(n_1, n_2)$ ，其概率密度为

$$f_F(z) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right)}{\Gamma\left(\dfrac{n_1}{2}\right)\Gamma\left(\dfrac{n_2}{2}\right)} n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}} \dfrac{z^{\frac{n_1}{2}-1}}{(n_1 z + n_2)^{\frac{n_1+n_2}{2}}} & z > 0 \\ 0 & z \le 0 \end{cases}$$

F分布

$F \sim F(n_1, n_2)$ n_1 为第一自由度；n_2 为第二自由度。



Remark 1：如果 $X \sim F(m, n)$，则 $\dfrac{1}{X} \sim F(n, m)$。

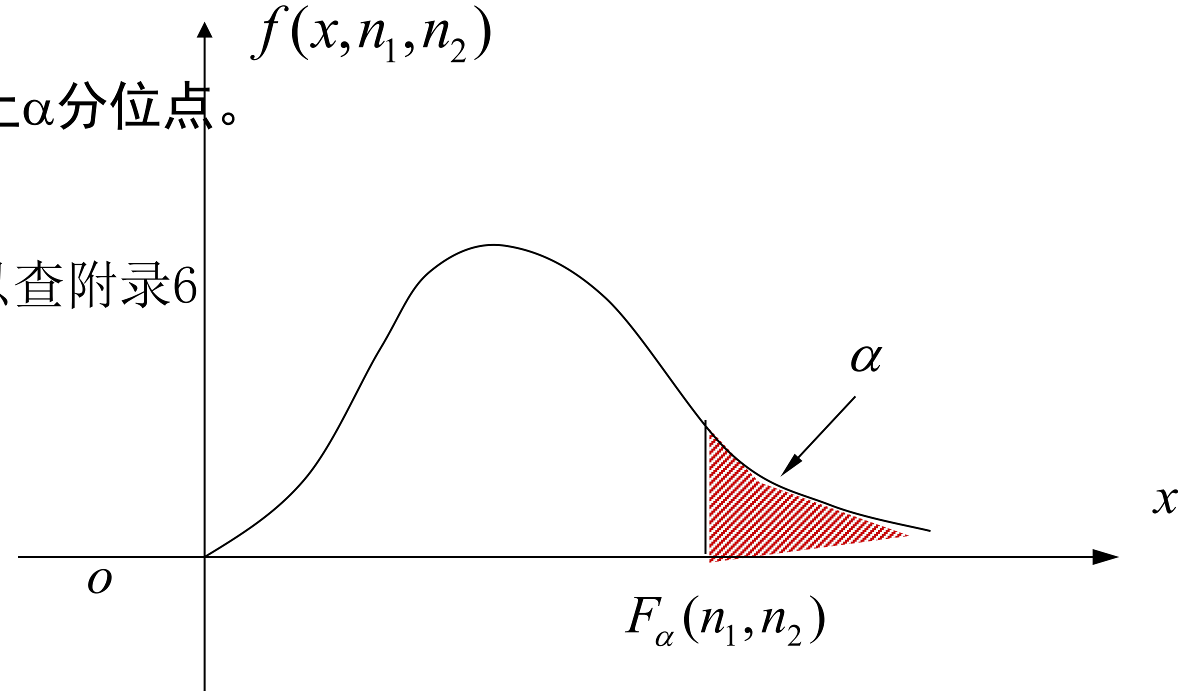
F分布

F 分布的分位点 对于给定的 $\alpha, 0<\alpha<1$，称满足条件

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

称 $F_\alpha(n_1, n_2)$ 为F分布的上α分位点。

注：F分布的上α分位点可以查附录6



Remark 2: $$F_\alpha(n_1, n_2) = \frac{1}{F_{1-\alpha}(n_2, n_1)}.$$ 思考：如何证明？

例题4：设X_1, X_2, X_3, X_4是来自总体N(0,4)的样本，

(1) 求常数C，使得Y=C[$(X_1 - X_2)^2 + (X_3 + X_4)^2$] 服从$\chi^2$分布，并指出自由度是多少？

(2)证明$Z = \frac{(X_1 - X_2)^2}{(X_3 + X_4)^2}$服从F(1,1)

[定理6.7] 设 $X_1, X_2, \cdots, X_{n_1}$ 与 $Y_1, Y_2, \cdots, Y_{n_2}$ 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，且两个样本独立。

$\overline{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i$，$\overline{Y} = \frac{1}{n_2}\sum_{j=1}^{n_2} Y_j$ 分别是两个样本的样本均值

$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_i - \overline{X})^2$，$S_2^2 = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}\left(Y_j - \overline{Y}\right)^2$ 是各样本方差

则有：

$$\frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

[1]

证明：

[1] 由 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, $X_1, X_2, \cdots, X_{n_1}$ 与 $Y_1, Y_2, \cdots, Y_{n_2}$ 分别是来自两个总体的简单样本

$X_1, X_2, \cdots, X_{n_1}$的线性组合 $\overline{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i$ 服从正态分布 $N(\mu_1, \frac{\sigma_1^2}{n_1})$

同理 $\overline{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n_2})$

又由于$\overline{X}$, $\overline{Y}$ 独立，其组合是正态分布，所以

$$\overline{X} - \overline{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

其标准化随机变量 $\dfrac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$

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

证明：

[2] $$\chi_1^2 = \frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1)$$

$$\chi_2^2 = \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2-1)$$

由于 S_1^2 与 S_2^2 独立，

$$\frac{\chi_1^2 / (n_1-1)}{\chi_2^2 / (n_2-1)} \sim F(n_1-1, n_2-1)$$

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

[3] 当$\sigma_1=\sigma_2=\sigma$时

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$$

其中 $S_w=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

证明：

[3] 随机变量 $U=\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim N(0,1)$，由上知，统计量

$$\chi_1^2=\frac{(n_1-1)S_1^2}{\sigma^2}\sim\chi^2(n_1-1)$$

$$\chi_2^2=\frac{(n_2-1)S_2^2}{\sigma^2}\sim\chi^2(n_2-1)$$

又由于 S_1^2 与 S_2^2 独立，利用χ^2 分布的可加性知随机变量

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

随机变量U,V独立, 因此

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

$$S_w=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

[定理6.8] 假设 $X_1,\cdots,X_n$ 是来自正态总体 $X\sim N(\mu,\sigma^2)$ 的独立样本，则样本均值 $\overline{X}$ 与样本方差 S^2 独立。

考研的试题中出现过本定理的应用！

例题5：设总体$X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$，从两个总体中分别抽样得到$n_1 = 8,\ S_1^2 = 8.75; n_2 = 10,\ S_2^2 = 2.66$。求概率$P\{\sigma_1^2 > \sigma_2^2\}$

例题6：设在总体$X\sim N(\mu,\sigma^2)$中抽样一容量为16的样本，这里μ,σ^2均为未知。

（1）求概率$P\left\{\frac{S^2}{\sigma^2}\le 2.041\right\}$,其中$S^2$为样本方差；（2）求$D(S^2)$

总结：几类抽样分布

- 样本均值分布 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,\ 1)$
- χ^2分布 $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2(n)$ $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
- T分布 $t=\dfrac{\overline{X}-\mu}{S\big/\sqrt{n}} \sim t(n-1)$
- F分布 $\dfrac{S_1^2\big/\sigma_1^2}{S_2^2\big/\sigma_2^2} = \dfrac{S_1^2\big/S_2^2}{\sigma_1^2\big/\sigma_2^2} \sim F(n_1-1, n_2-1)$

注意：以上是对X~N(μ，σ²）而言的；取得 n 个样本