概率论与数理统计

第 17 讲

抽样分布

一. 三大重要分布

统计量是样本$X_1, X_2, \ldots, X_n$的函数，而样本是随机变量。因此，统计量也是随机变量，**统计量的分布称为抽样分布**。

近代统计学的创始人之一，英国统计学家费歇(R.A.Fisher)把抽样分布、参数估计和假设检验列为统计推断的三个中心内容.

本节给出数理统计中的三大分布：χ^2分布，t分布，F分布及其几个重要的抽样分布定理。

定义1 设 $X_1, X_2, \dots, X_n$ 是来自总体 $N(0, 1)$的样本，则称统计量

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

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

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

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

其中：伽玛函数$\Gamma(\alpha)$通过下述积分定义

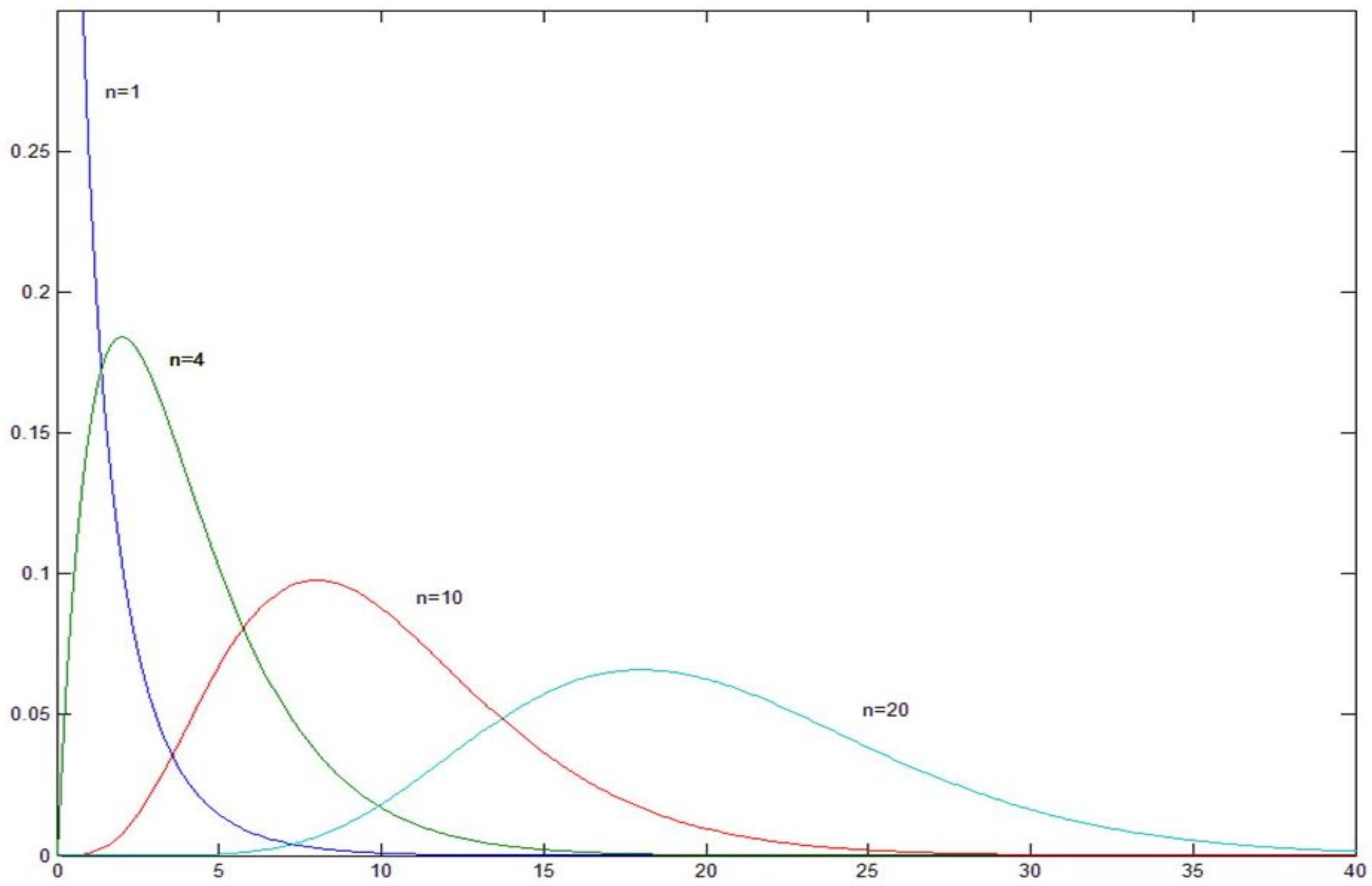
$$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}e^{-x}dx,\quad \alpha>0$$

特别地，当$n=2$时，$\chi^2(2)$分布的概率密度为

期望为2的指数分布

$$f(x)=\begin{cases}\dfrac{1}{2}e^{-\frac{x}{2}}, & x>0\\ 0, & x\le 0\end{cases}$$

一. 三大重要分布—χ^2分布



χ^2 分布的性质

1. 设 $X_1, X_2, \ldots, X_n$ 相互独立，都服从正态分布 $N(\mu, \sigma^2)$，则

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

证明：由 $X_i \sim N(\mu, \sigma^2)$ 且独立，得 $\dfrac{X_i-\mu}{\sigma} \sim N(0,1)$ 且相互独立，于是

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

一. 三大重要分布—χ^2分布

2. 设 $X\sim\chi^2(n)$，则 $E(X)=n$，$D(X)=2n$。

证明：易知存在 $X_i\sim N(0,1)$，$i=1, 2,\ldots, n$ 且相互独立，使得 $X=\sum_{i=1}^{n}X_i^2$，于是

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

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

$$E(X_i^2)=D(X_i)+[E(X_1)]^2=1$$

$$D(X_i^2)=E[(X_i^2)^2]-[E(X_i^2)]^2$$
$$=3-1=2$$

$$E(X_i^4)=\int_{-\infty}^{+\infty}x^4\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$
$$=3$$

3. 设 $X_1\sim\chi^2(n_1)$, $X_2\sim\chi^2(n_2)$，且 X_1, X_2 相互独立，则

$$X_1+X_2\sim\chi^2(n_1+n_2)$$

证明：由 χ^2分布的定义知，存在$U_i\sim N(0,1)$, i=1, 2 ,…, n_1； $V_j\sim N(0,1)$, j=1, 2 ,…, n_2；所有变量相互独立，满足 $X_1=U_1^2+U_2^2+\cdots+U_{n_1}^2$， $X_2=V_1^2+V_2^2+\cdots+V_{n_2}^2$ ，于是有

$$X_1+X_2=U_1^2+U_2^2+\cdots+U_{n_1}^2+V_1^2+V_2^2+\cdots+V_{n_2}^2\sim\chi^2(n_1+n_2)$$

推广：设 $X_i\sim\chi^2(n_i)$, i=1,2,…,k，且相互独立，则

$$\sum_{i=1}^{k}X_i\sim\chi^2(\sum_{i=1}^{k}n_i)$$

χ^2 分布的可加性

例1. 设总体 $X \sim N(0,1)$, $X_1, X_2, \ldots, X_6$ 为来自总体 X 的样本, 记 $Y=(X_1+X_2+X_3)^2+(X_4+X_5+X_6)^2$，试确定常数 c，使 cY 服从χ^2分布。

解：因为 $X_1+X_2+X_3 \sim N(0,3)$， $X_4+X_5+X_6 \sim N(0,3)$，所以有

$$(X_1+X_2+X_3)\big/\sqrt{3} \sim N(0,1),\quad (X_4+X_5+X_6)\big/\sqrt{3} \sim N(0,1)$$

且上述两个标准正态随机变量相互独立，故

$$\left(\frac{X_1+X_2+X_3}{\sqrt{3}}\right)^2+\left(\frac{X_4+X_5+X_6}{\sqrt{3}}\right)^2=\frac{Y}{3} \sim \chi^2(2)$$

因此 $c=1/3$.

例2. 设总体 X 服从指数分布$E(\lambda)$，$X_1, X_2, \ldots, X_n$是来自总体 X 的样本，证明

$$T=2\lambda(X_1+X_2+\cdots+X_n)=2n\lambda\bar{X}\sim\chi^2(2n)$$

证明：易知 $X_1, X_2, \ldots, X_n$独立同分布，都服从$E(\lambda)$，其密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$$

易求得$Y_1=2\lambda X_1$的密度函数 $f_{Y_1}(y)=\begin{cases}\dfrac{1}{2}e^{-\frac{1}{2}x}, & x>0\\ 0, & x\le 0\end{cases}$**，恰好是自由度为2的 χ^2 分布。**

于是有 $2\lambda X_1, 2\lambda X_2, \ldots, 2\lambda X_n$ 独立同分布且都服从自由度为 2 的 χ^2 分布，根据 χ^2 分布布的可加性得

$$2\lambda X_1 + 2\lambda X_2 + \cdots + 2\lambda X_n \sim \chi^2(2n)$$

即有，

$$T = 2\lambda(X_1 + X_2 + \cdots + X_n) = 2n\lambda \bar{X} \sim \chi^2(2n)$$

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

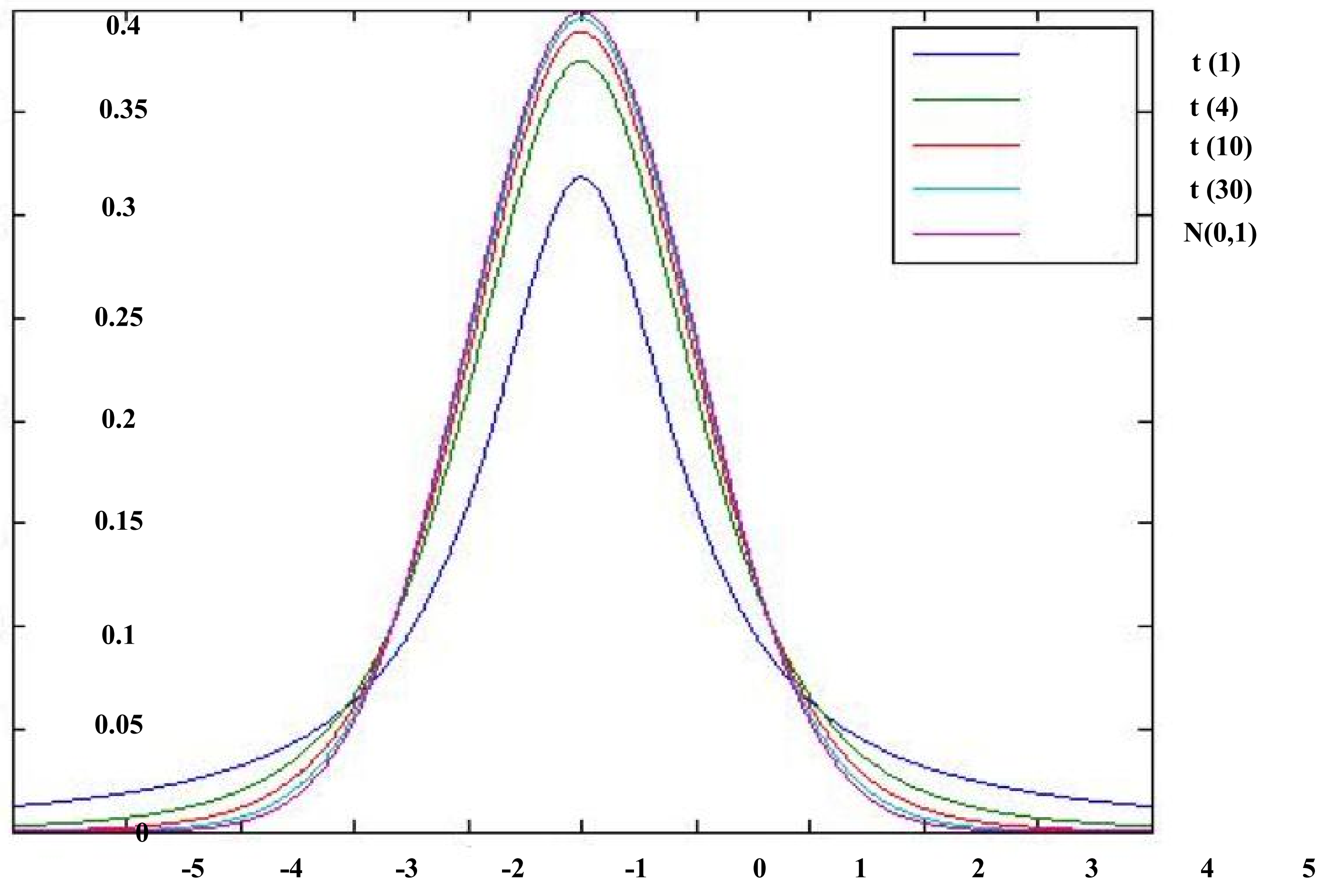
$$t=\frac{X}{\sqrt{Y/n}}$$

服从自由度为n的 t 分布，又称学生氏（*student*）分布，记为$t\sim t(n)$，其概率密度函数为

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

一. 三大重要分布—t 分布

$t(n)$分布的密度函数



t 分布的性质

1. t 分布的密度函数关于 y 轴对称，且 $\lim\limits_{|x|\to\infty} f(x)=0$

2. t 分布的密度函数形状是中间高，两边低，左右对称，与标准正态分布的概率密度函数图像类似。

---当 n 较大时，T 近似服从标准正态分布$N(0,1)$；

---当 n 较小时，在尾部 t 分布比标准正态分布有更大的概率。

一. 三大重要分布—t 分布

3. 设$T\sim t(n)$，$n>1$，则对于$r<n$，$E(T^r)$ 存在，且

$$E(T^r)=\begin{cases} n^{\frac{r}{2}}\dfrac{\Gamma[(r+1)/2]\Gamma[(n-r)/2]}{\sqrt{\pi}\Gamma(n/2)}, & \text{当}r\text{是偶数}, \\ 0, & \text{当}r\text{是奇数}. \end{cases}$$

特别的，$E(T)=0$，$D(T)=n/(n-2)$，对于$n>2$。

4. 当$n=1$时，其密度函数为

$$f(x)=\frac{1}{\pi(1+x^2)},\quad -\infty<x<\infty$$

此时，t 分布就是柯西分布，其数学期望不存在。

例3. 设 X 与 Y 相互独立, $X\sim N(0,16)$, $Y\sim N(0,9)$, $X_1, X_2,\dots, X_9$ 与 $Y_1, Y_2,\dots, Y_{16}$ 分别是取自 X 与 Y 的简单随机样本。求统计量

$$\frac{X_1+X_2+\cdots+X_9}{\sqrt{Y_1^2+Y_2^2+\cdots+Y_{16}^2}}$$

所服从的分布。

解：因为 $X_1+X_2+\ldots+X_9\sim N(0,144)$，所以有

$$\frac{1}{12}(X_1+X_2+\cdots+X_9)\sim N(0,1)$$

另外，因为

$$\frac{1}{3}Y_i\sim N(0,1),\ i=1,2,\ldots 16$$

$$\frac{X_1+X_2+\cdots+X_9}{\sqrt{Y_1^2+Y_2^2+\cdots+Y_{16}^2}},\qquad \frac{1}{3}Y_i\sim N(0,1),\ i=1,2,\ldots 16,$$

$$\frac{1}{12}(X_1+X_2+\cdots+X_9)\sim N(0,1)$$

所以

$$\sum_{i=1}^{16}\left(\frac{1}{3}Y_i\right)^2\sim\chi^2(16)$$

且上述两个随机变量相互独立，因此

$$\frac{(X_1+X_2+\cdots+X_9)/12}{\sqrt{\frac{1}{16}\sum_{i=1}^{16}\left(\frac{1}{3}Y_i\right)^2}}=\frac{X_1+X_2+\cdots+X_9}{\sqrt{Y_1^2+Y_2^2+\cdots+Y_{16}^2}}\sim t(16)$$

一. 三大重要分布—F 分布

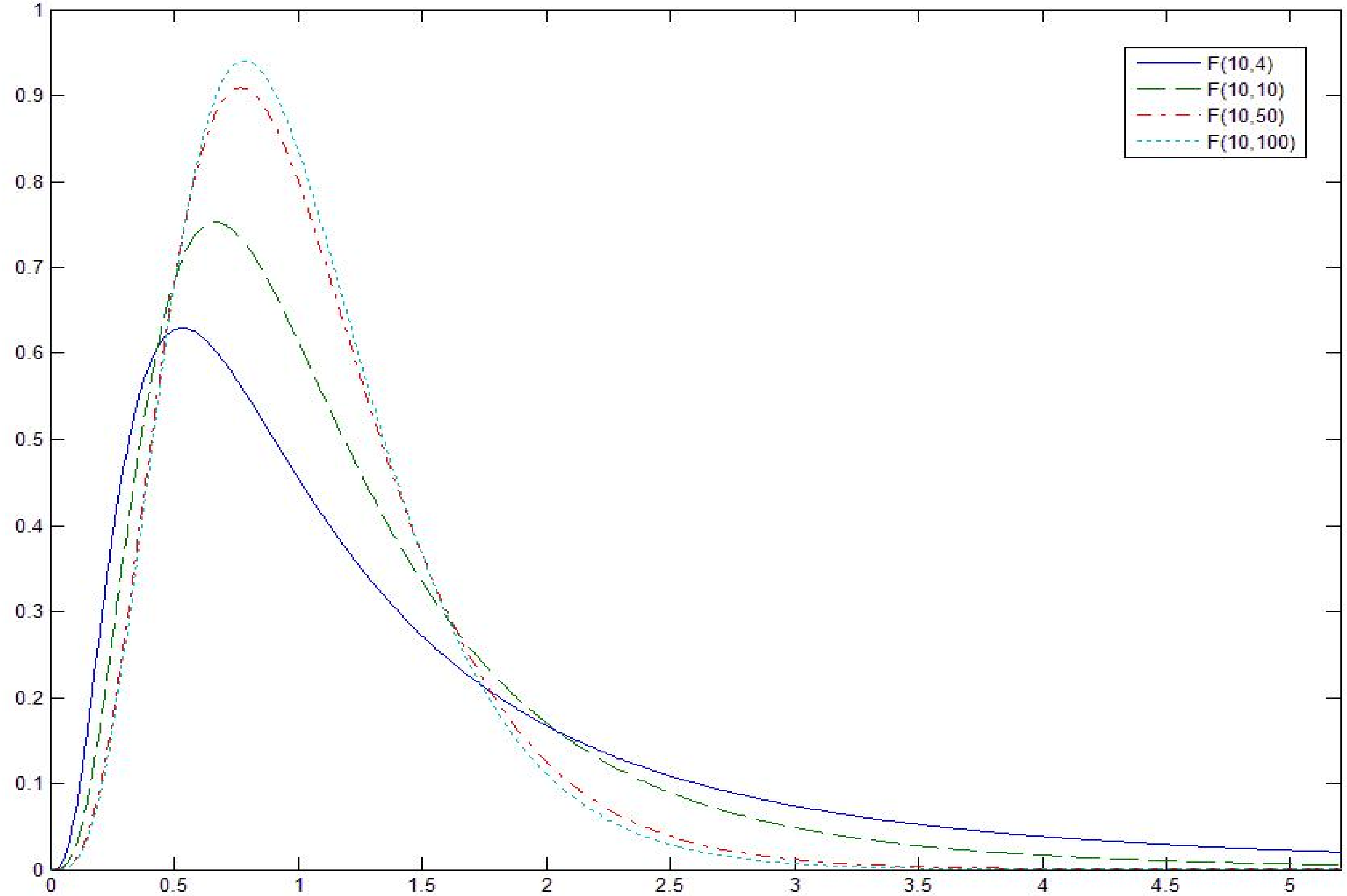
定义3 设 $X\sim \chi^2(m)$, $Y\sim\chi^2(n)$，且 X 与 Y 相互独立，称随机变量

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

服从自由度为$(m,\ n)$的 F 分布，记为$F\sim F(m,n)$，其概率密度函数为

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

一. 三大重要分布—F 分布


F(10,4)
F(10,10)
F(10,50)
F(10,100)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

F 分布的性质

1. 设 $X\sim F(m, n)$，则$1/X \sim F(n, m)$

2. 设 $T\sim t(n)$，则$T^2 \sim F(1, n)$

证明：由于$T \sim t(n)$，所以

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

其中$X\sim N(0,1)$, $Y\sim \chi^2(n)$ 且 X 与 Y 独立，于是有

$$T^2=\left(\frac{X}{\sqrt{Y/n}}\right)^2=\frac{X^2}{Y/n} \sim F(1, n)$$

例4. 设随机变量 X 和 Y 都服从标准正态分布, 则

(1) $X+Y$ 服从正态分布　　(2) X^2+Y^2 服从 χ^2 分布

(3) X^2 和 Y^2 都服从 χ^2 分布　√

(4) X^2/Y^2 服从 F 分布

引理： 设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的一个样本，且 $EX=\mu$, $DX=\sigma^2$，则有

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

其中 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 为样本均值，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 为样本方差。

证明：易知 $X_1, X_2, \ldots, X_n$ 独立同分布，且期望均为 μ，方差均为 σ^2。于是由

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

二. 几个重要的抽样分布

$$D(\bar{X})=D(\frac{1}{n}\sum_{i=1}^{n}X_i)=\frac{1}{n^2}\sum_{i=1}^{n}D(X_i)=\frac{\sigma^2}{n}$$

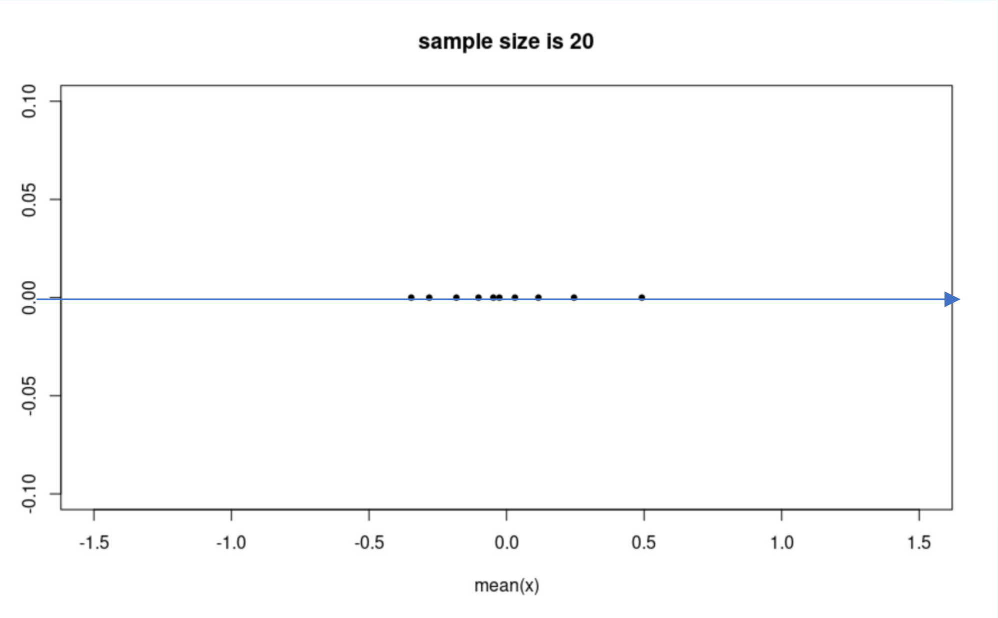
$$E(S^2)=E[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2]=\frac{1}{n-1}E[\sum_{i=1}^{n}(X_i-\mu+\mu-\bar{X})^2]$$

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

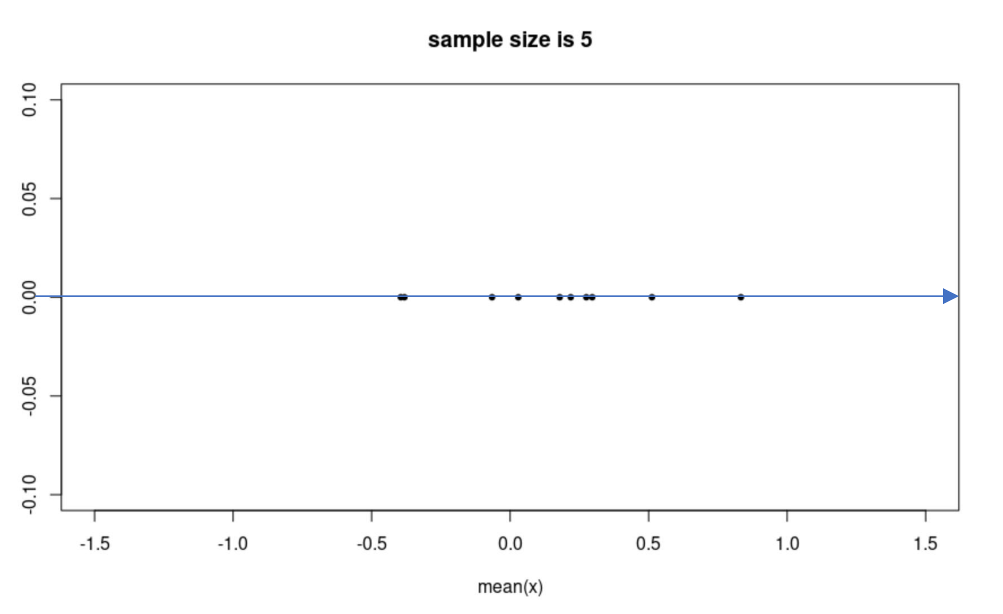
$$=\frac{1}{n-1}[\sum_{i=1}^{n}E(X_i-\mu)^2-nE(\bar{X}-\mu)^2]$$

$$=\frac{1}{n-1}[\sum_{i=1}^{n}DX_i-nD\bar{X}]=\frac{1}{n-1}(\sum_{i=1}^{n}\sigma^2-n\frac{\sigma^2}{n})=\sigma^2$$

二. 几个重要的抽样分布



$$E\bar{X}=\mu\,,\,D\bar{X}=\frac{\sigma^2}{n}$$



二. 几个重要的抽样分布

定理1. 设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$, $\mu \in R, \sigma > 0$ 的样本，$\bar{X}$ 表示样本均值，则有

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

证明：易知 $X_1, X_2, \ldots, X_n$ 服从 $N(\mu, \sigma^2)$ 且相互独立。由于相互独立正态随机变量的线性组合仍然服从正态分布，而且 $E\bar{X} = \mu$, $D\bar{X} = \sigma^2 / n$. 所以有

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

将其标准化得

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

定理2. 设 $X_1, X_2, \dots, X_n$是来自正态总体$N(\mu, \sigma^2), \mu \in R, \sigma > 0$的样本，$\bar{X}$ 和 S^2 分别表示样本均值和样本方差，则有

（1）$\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$;　（2）$\bar{X}$ 和 S^2相互独立

定理3. 设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$, $\mu \in R$, $\sigma > 0$ 的样本，$\bar{X}$ 和 S^2 分别表示样本均值和样本方差，则有

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

证明：由定理 1 和定理 2，得

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

且它们相互独立，根据 t 分布的定义得

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \Bigg/ \sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$

定理4. 设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N(\mu_1,\sigma_1^2)$ 的样本，设 $Y_1, Y_2, \ldots, Y_m$ 是来自正态总体 $N(\mu_2, \sigma_2^2)$ 的样本，且两个样本相互独立，$\bar{X},\bar{Y}$ 分别是这两个样本的样本均值，则有

$$Z=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}\sim N(0,1)$$

证明：由定理1得

$$\bar{X}\sim N(\mu_1,\frac{\sigma_1^2}{n}),\bar{Y}\sim N(\mu_2,\frac{\sigma_2^2}{m})$$

二. 几个重要的抽样分布

$$\bar{X} \sim N(\mu_1, \frac{\sigma^2}{n}), \bar{Y} \sim N(\mu_2, \frac{\sigma^2}{m})$$

又由正态分布性质得

$$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m})$$

标准化后有

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim N(0,1)$$

定理5. 设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N(\mu_1,\sigma^2)$ 的样本，设 $Y_1, Y_2, \ldots, Y_m$ 是来自正态总体 $N(\mu_2, \sigma^2)$ 的样本，且两个样本相互独立，$\bar{X},\bar{Y},S_1^2,S_2^2$ 分别是这两个样本的样本均值和样本方差。则有

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_\omega\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}} \sim t(n+m-2)$$

其中

$$S_\omega^2=\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2},\quad S_\omega=\sqrt{S_\omega^2}$$

证明：由定理1得

$$\bar{X} \sim N(\mu_1, \frac{\sigma^2}{n}), \bar{Y} \sim N(\mu_2, \frac{\sigma^2}{m})$$

又由正态分布性质得

$$\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma^2}{n} + \frac{\sigma^2}{m})$$

标准化后有

$$\boxed{U = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0,1)}$$

根据定理 2 有

$$\frac{(n-1)S_1^2}{\sigma^2} \sim \chi^2(n-1), \qquad \frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2(m-1)$$

且它们相互独立，又由 χ^2 分布的可加性得到

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

二. 几个重要的抽样分布

$$U=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sigma\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}\sim N(0,1)$$

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

$$S_\omega^2=\frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2},\quad S_\omega=\sqrt{S_\omega^2}$$

易知U和V是互相独立性的，根据t分布的定义得到

$$\frac{U}{\sqrt{V/(n+m-2)}}=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_\omega\sqrt{1/n+1/m}}\sim t(n+m-2)$$

二. 几个重要的抽样分布

定理6. 设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N(\mu_1,\sigma_1^2)$ 的样本，设$Y_1, Y_2, \ldots, Y_m$ 是来自正态总体 $N(\mu_2,\sigma_2^2)$ 的样本，且两个样本相互独立，S_1^2, S_2^2 分别是这两个样本样本方差。则有 $\dfrac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n-1, m-1)$

证明：根据定理 2 有

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

根据 F 分布的定义得

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

例5. 设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $N(\mu,\sigma^2)$, $\mu\in R,\sigma>0$ 的样本, $\bar{X}$ 和 S^2 分别表示样本均值和样本方差，又设 $X_{n+1}\sim N(\mu,\sigma^2)$，且 X_{n+1} 与 $X_1, X_2, \dots, X_n$ 互相独立，证明

$$T=\sqrt{\frac{n}{n+1}}\frac{\bar{X}-X_{n+1}}{S}\sim t(n-1)$$

证明：易知

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

且它们相互独立。由于 X_{n+1} 与 $X_1, X_2, \dots, X_n$ 互相独立，所以 $\bar{X}$ 与 X_{n+1} 互相独立，

$$E(\bar{X}-X_{n+1})=E(\bar{X})-E(X_{n+1})=\mu-\mu=0$$

二. 几个重要的抽样分布

$$E(\bar{X}-X_{n+1})=E(\bar{X})-E(X_{n+1})=\mu-\mu=0$$

$$D(\bar{X}-X_{n+1})=D(\bar{X})+D(X_{n+1})=\frac{1}{n}\sigma^2+\sigma^2=\frac{n+1}{n}\sigma^2$$

所以

$$\bar{X}-X_{n+1}\sim N(0,\frac{n+1}{n}\sigma^2)$$

从而

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

二. 几个重要的抽样分布

$$\frac{\bar{X}-X_{n+1}}{\sqrt{\frac{n+1}{n}}\sigma} \sim N(0,1) \quad , \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

因为 $\bar{X}-X_{n+1}$ 与 $\frac{(n-1)S^2}{\sigma^2}$ 相互独立。由 t 分布的定义得

$$T=\frac{\bar{X}-X_{n+1}}{\sqrt{\frac{n+1}{n}}\sigma} \Bigg/ \sqrt{\frac{(n-1)S^2}{(n-1)\sigma^2}}=\sqrt{\frac{n}{n+1}}\frac{\bar{X}-X_{n+1}}{S} \sim t(n-1)$$

例6. 设总体 $X\sim N(1,\sigma^2)$，总体 $Y\sim N(2,\sigma^2)$，且 X 与 Y 相互独立，$X_1,X_2,\dots,X_m$ 是来自总体 X 的样本，$Y_1,Y_2,\dots,Y_n$ 是来自总体 Y 的样本。$\bar{X},\bar{Y},S_1^2,S_2^2$ 分别是这两个样本的样本均值和样本方差。α 和 β 是两个固定的实数，求 T 的分布。

$$T=\frac{\alpha(\bar{X}-1)+\beta(\bar{Y}-2)}{\sqrt{\dfrac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}}\sqrt{\dfrac{\alpha^2}{m}+\dfrac{\beta^2}{n}}}$$

解：由于 $\bar{X}$, $\bar{Y}$ 相互独立，且

$$\bar{X}\sim N(1,\frac{\sigma^2}{m}),\bar{Y}\sim N(2,\frac{\sigma^2}{n})$$

$$\bar{X} \sim N(1, \frac{\sigma^2}{m}), \bar{Y} \sim N(2, \frac{\sigma^2}{n})$$

因为

$$E[\alpha(\bar{X}-1)+\beta(\bar{Y}-2)] = E[\alpha(\bar{X}-1)] + E[\beta(\bar{Y}-2)]$$

$$= \alpha E(\bar{X}-1) + \beta E(\bar{Y}-2) \quad = 0$$

$$D[\alpha(\bar{X}-1)+\beta(\bar{Y}-2)] = D[\alpha(\bar{X}-1)] + D[\beta(\bar{Y}-2)]$$

$$= \alpha^2 D(\bar{X}-1) + \beta^2 D(\bar{Y}-2) \quad = \alpha^2 D(\bar{X}) + \beta^2 D(\bar{Y})$$

$$= \alpha^2 \frac{\sigma^2}{m} + \beta^2 \frac{\sigma^2}{n} \quad = (\frac{\alpha^2}{m} + \frac{\beta^2}{n})\sigma^2$$

$$E[\alpha(\bar{X}-1)+\beta(\bar{Y}-2)]=0,$$

$$D[\alpha(\bar{X}-1)+\beta(\bar{Y}-2)]=(\frac{\alpha^2}{m}+\frac{\beta^2}{n})\sigma^2$$

根据相互独立正态随机变量线性组合仍然服从正态分布得到

$$U=\frac{\alpha(\bar{X}-1)+\beta(\bar{Y}-2)}{\sigma\sqrt{\frac{\alpha^2}{m}+\frac{\beta^2}{n}}}\sim N(0,1)$$

又有

$$\frac{(m-1)S_1^2}{\sigma^2}\sim\chi^2(m-1),\qquad \frac{(n-1)S_2^2}{\sigma^2}\sim\chi^2(n-1)$$

所以

$$\frac{(m-1)S_1^2+(n-1)S_2^2}{\sigma^2}\sim\chi^2(n+m-2)$$

二. 几个重要的抽样分布

$$U=\frac{\alpha(\bar{X}-1)+\beta(\bar{Y}-2)}{\sigma\sqrt{\dfrac{\alpha^2}{m}+\dfrac{\beta^2}{n}}}\sim N(0,1),$$

$$\frac{(m-1)S_1^2+(n-1)S_2^2}{\sigma^2}\sim\chi^2(n+m-2)$$

易知 U 与 V 相互独立。因此，根据 t 分布的定义有

$$\frac{U}{\sqrt{V/(m+n-2)}}=\frac{\alpha(\bar{X}-1)+\beta(\bar{Y}-2)}{\sqrt{\dfrac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}}\sqrt{\dfrac{\alpha^2}{m}+\dfrac{\beta^2}{n}}}$$

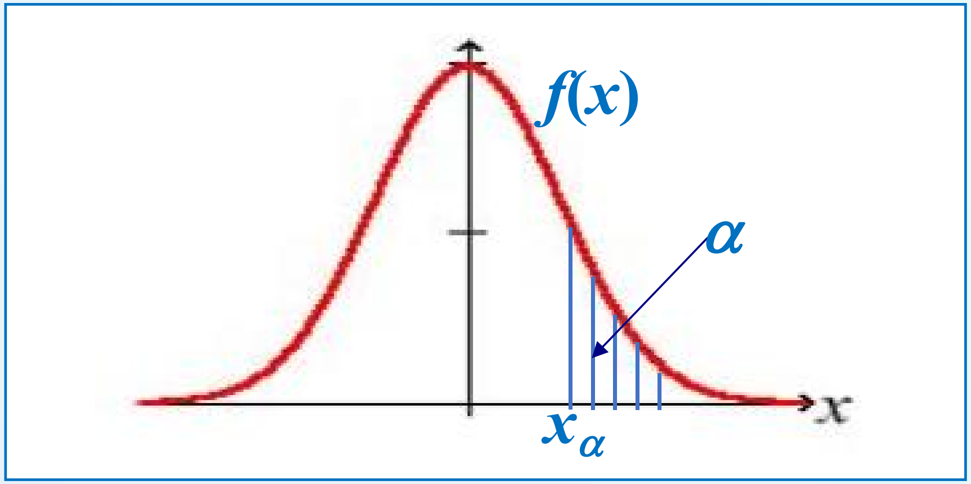
$$\sim t(m+n-2)$$

三. 分位数

定义4. 设 $0<\alpha<1$，对连续型随机变量 X，称满足 $P\{X>x_\alpha\}=\alpha$ 的点 x_α 为 X 的概率分布的上 α 分位数(分位点)。

若 X 的概率密度函数 $f(x)$，则 x_α 满足

$$P\{X>x_\alpha\}=\int_{x_\alpha}^{+\infty}f(x)dx=\alpha$$

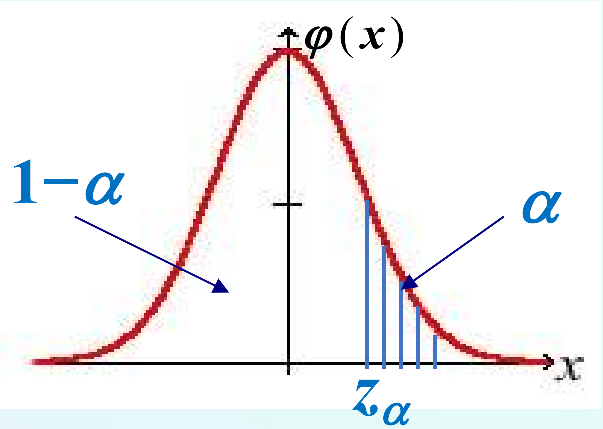


三. 分位数

$$P\{X > x_\alpha\} = \int_{x_\alpha}^{+\infty} f(x)dx = \alpha$$

定义2. 设随机变量 $X \sim N(0,1)$，对 $0<\alpha<1$，称满足 $P\{X>z_\alpha\}=\alpha$ 的点 z_α 为标准正态分布的上 α 分位数(分位点)。

性质1. $\Phi(z_\alpha) = P(X \le z_\alpha)=1-\alpha$



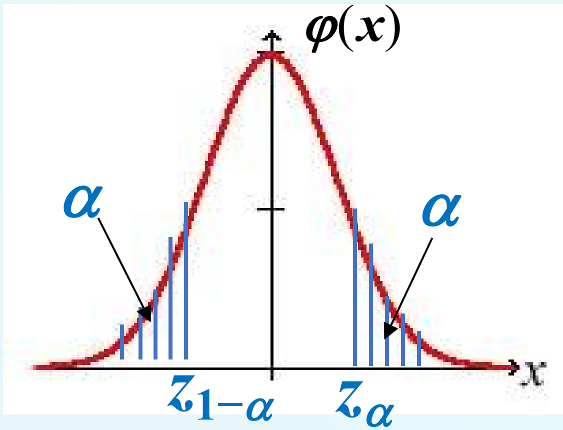
三. 分位数

$$P\{X > x_\alpha\} = \int_{x_\alpha}^{+\infty} f(x)dx = \alpha \qquad P\{X \le x_\alpha\} = 1-\alpha$$

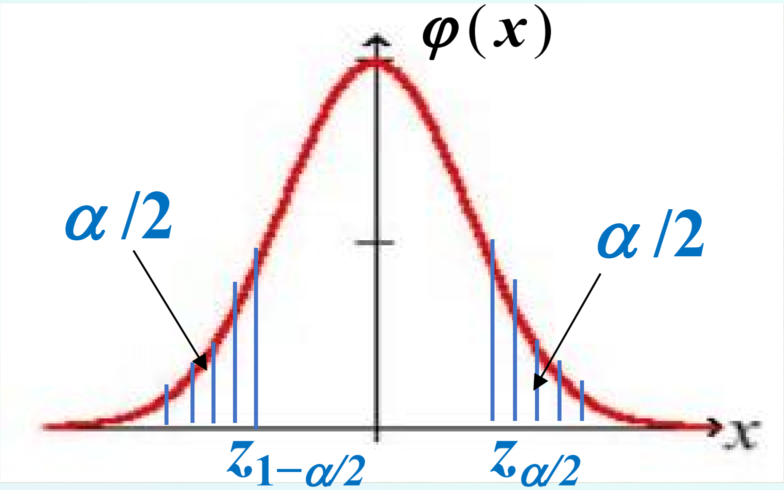
性质2. $z_{1-\alpha} = -z_\alpha$

常用数字 $z_{0.025} = 1.96,\quad z_{0.975} = -1.96,\quad z_{0.05} = 1.645,$

$z_{0.95} = -1.645.$



三. 分位数



两个常用式子

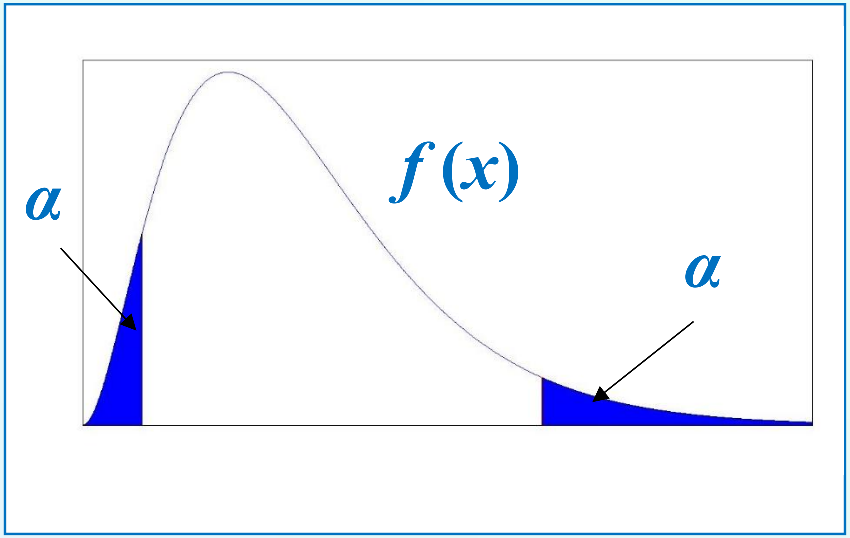
$$P\{|Z|>z_{\alpha/2}\}=\alpha,\quad P\{|Z|\le z_{\alpha/2}\}=1-\alpha$$

定义5. 设随机变量 $X\sim \chi^2(n)$，对 $0<\alpha<1$，称满足 $P\{X>\chi^2_\alpha(n)\}=\alpha$ 的点 $\chi^2_\alpha(n)$ 为自由度为 n 的 χ^2 分布的上 α 分位数(分位点)。

如： $\chi^2_{0.05}(6)=12.592$

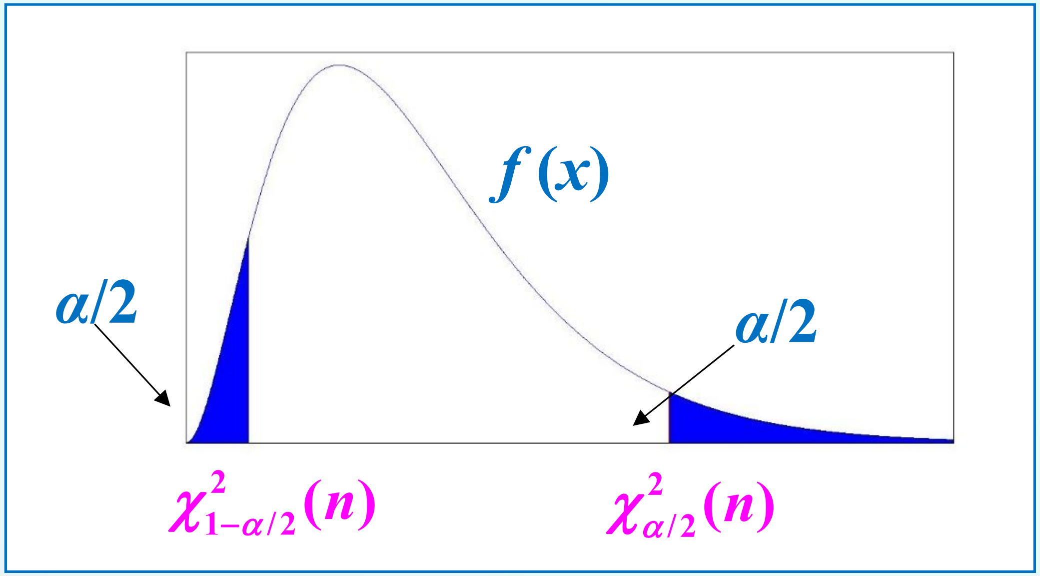
当n充分大($n>45$)时

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



其中 z_α 是标准正态分布的上 α 分位点。

三. 分位数



常用两个式子

$$P\{\chi^2 > \chi^2_{\alpha/2}(n)\} + P\{\chi^2 < \chi^2_{1-\alpha/2}(n)\} = \alpha$$

$$P\{\chi^2_{1-\alpha/2}(n) \le \chi^2 \le \chi^2_{\alpha/2}(n)\} = 1-\alpha$$

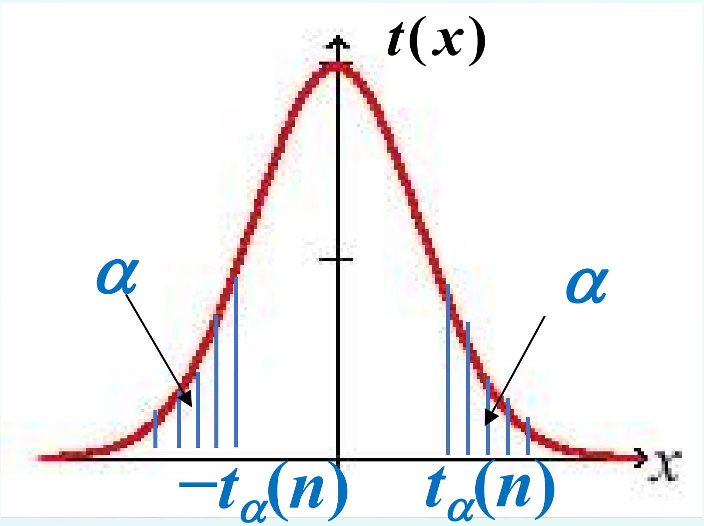
三. 分位数

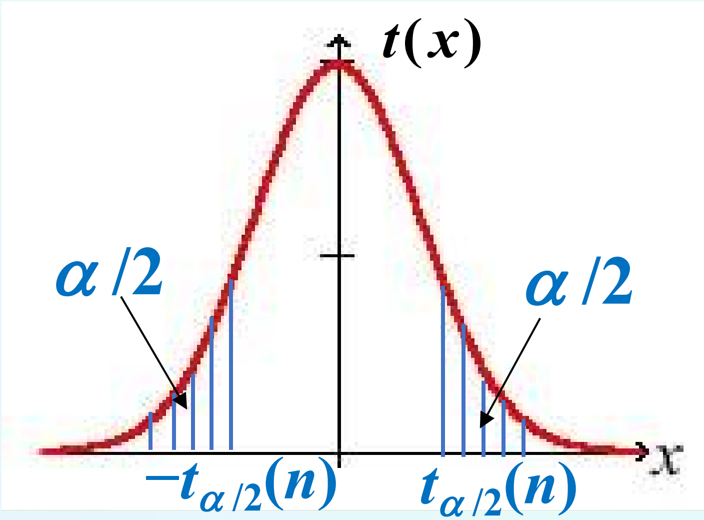
定义6. 设随机变量 $X\sim t(n)$，对$0<\alpha<1$，称满足 $P\{X>t_\alpha(n)\}=\alpha$ 的点$t_\alpha(n)$为自由度为 n 的 t 分布的上 α 分位数(分位点)。

性质1. $t_{1-\alpha}(n)=-t_\alpha(n)$

性质2. 当 n 较大($n>45$)时，

$$t_\alpha(n)\approx z_\alpha$$





性质2. 若 $P(|X|>c) = P\{(X>c)\cup(X<-c)\}$

$$= P(X>c)+P(X<-c)\}=\alpha,$$

则 $c = t_{\alpha/2}(n)$

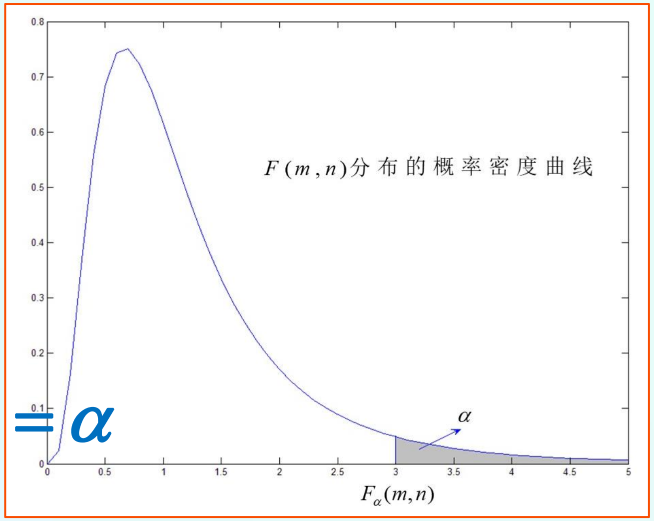
两个常用式子：

$$P\{|t|>t_{\alpha/2}(n)\}=\alpha,\quad P\{|t|\le t_{\alpha/2}(n)\}=1-\alpha$$

三. 分位数

定义7. 设随机变量 $X\sim F(m,n)$，对 $0<\alpha<1$，称满足 $P\{X>F_\alpha(m,n)\}=\alpha$ 的点 $F_\alpha(m,n)$ 为自由度为 m,n 的 F 分布的上α 分位数(分位点)。

如：$F_{0.05}(6,9)=3.37$



两个常用式子：

$$P\{F>F_{\alpha/2}(m,n)\}+P\{F<F_{1-\alpha/2}(m,n)\}=\alpha$$

$$P\{F_{1-\alpha/2}(m,n)\le F\le F_{\alpha/2}(m,n)\}=1-\alpha$$

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

证明：因为 $X\sim F(m,n)$，**所以** $1/X\sim F(n,m)$，

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

由此有

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

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

于是

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

作业： 7, 9, 10, 11, 12, 16, 19

第 17 讲

谢谢观看