§7.2 常用统计分布

(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：

$1^0.\chi_1^2\sim\chi^2(n_1),\chi_2^2\sim\chi^2(n_2)$,且$\chi_1^2$，$\chi_2^2$独立，则有

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

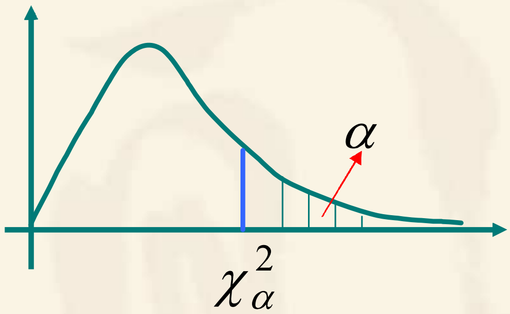
$2^0 . E\chi^2 = n, \quad D\chi^2 = 2n$

证：$EX_i = 0, \quad DX_i = 1, \quad X_i \sim N(0,1) \quad EX_i^2 = 1,$

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

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

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



对于给定的$\alpha(0<\alpha<1)$，称满足条件：

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

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

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

z_α是标准正态分布的上α分位点。

(2) $t-$分布

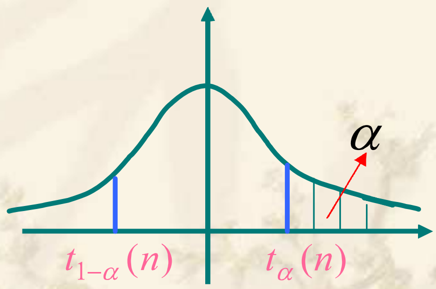
$X\sim N(0,1), Y\sim \chi^2(n), X,Y$独立，则称随机变量 $T=\dfrac{X}{\sqrt{Y/n}}$

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，$t_\alpha(n)\approx z_\alpha$.

(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

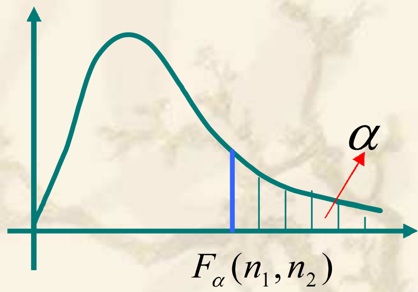
$$F=\frac{X/n_1}{Y/n_2}$$ 所服从的分布为自由度

是n_1,n_2 的 $F-$分布，记作$F\sim F(n_1,n_2)$.

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

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：$F_{0.95}(12,9) = \frac{1}{F_{0.05}(9,12)} = \frac{1}{2.80} = 0.357$

(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，$\overline{X},S^2$分别是样本均值与样本方差，则有：

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

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

(3). $\overline{X}$与S^2独立。

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

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

且它们独立。

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

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

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$\bar{X}=\dfrac{1}{n_1}\sum_{i=1}^{n_1}X_i,\bar{Y}=\dfrac{1}{n_2}\sum_{j=1}^{n_2}Y_j$ 分别是两个样本的均值。$S_1^2=\dfrac{1}{n_1-1}\sum_{i=1}^{n_1}(X_i-\bar{X})^2$

$S_2^2=\dfrac{1}{n_2-1}\sum_{j=1}^{n_2}(Y_j-\bar{Y})^2$分别是两个样本的方差；

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

证：$\bar{X}-\bar{Y} \sim N(\mu_1-\mu_2, \frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2})$

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

$\frac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2(n_1-1), \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_2-1)$,且它们独立。

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

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