第六章 数理统计的基本概念

第一节 总体与随机样本

第二节 常用统计量

第三节 常用统计量的分布---抽样分布

教学计划：1次课-3学时

概率论研究解决的问题

已知随机变量概率分布，$\begin{cases}\text{类型已知，但参数未知；} & \surd \\ \text{类型未知} & \times\end{cases}$

⇧

数理统计方法

$X \sim B(1,p)$

$X \sim B(n,p)$

$X \sim P(\lambda)$

$X \sim U(a,b)$

$X \sim E(\theta)$

$X \sim N(\mu,\sigma^2)$

解决两大问题：

1. 算概率：

$$P(x_1 < X \le x_2) = \sum_{x_1 < x_k \le x_2} p_k = \int_{x_1}^{x_2} f(t)\mathrm{d}t$$

$$P\{(X,Y)\in G\} = \sum_{(x_i,y_j)\in G} p_{ij} = \iint_G f(x,y)\mathrm{d}x\mathrm{d}y$$

2. 算数字特征：$E(X) = \int_{-\infty}^{+\infty} x\cdot f(x)\,\mathrm{d}x$

$$E[g(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$$

数理统计研究解决的问题

$$E(X)=\int_{-\infty}^{+\infty} x f(x)\mathrm{d}x$$

$X \sim B(1,p)$
$X \sim B(n,p)$
$X \sim P(\lambda)$
$X \sim U(a,b)$
$X \sim E(\theta)$
$X \sim N(\mu,\sigma^2)$

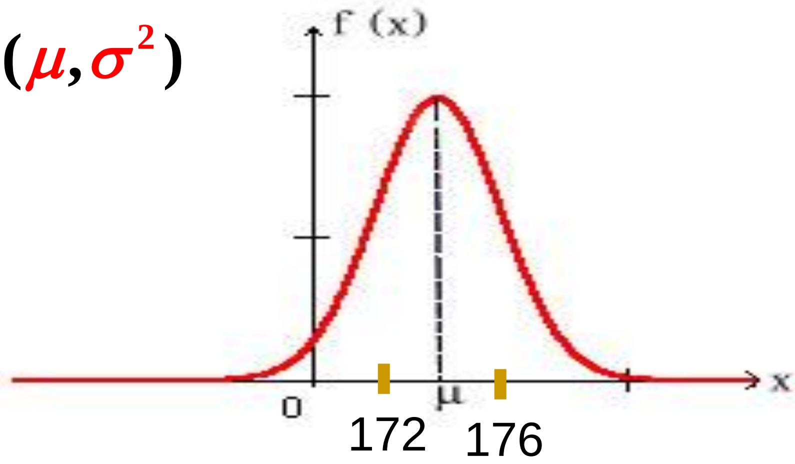
实际问题中，往往根据经验可知X 概率分布类型，但其中有未知参数。

设 $X \sim F(x,\theta)$, θ 是未知参数

概率分布：分布律　概率密度

例如：男生的身高X服从正态分布 $N(\mu,\sigma^2)$

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



$\mu=E(X)$, $\sigma^2=D(X)$---未知参数

---统计方法

$P(172\le X\le 176)$

$$=P\left(\frac{172-\mu}{\sigma}\le\frac{X-\mu}{\sigma}\le\frac{176-\mu}{\sigma}\right)=\Phi\left(\frac{176-\mu}{\sigma}\right)-\Phi\left(\frac{172-\mu}{\sigma}\right)$$

第六章 数理统计的基本概念

第一节 总体与随机样本

第二节 常用统计量

第三节 常用统计量的分布---抽样分布

教学计划：1次课-3学时

第六章 数理统计的基本概念

第一节 总体与随机样本

- 总体与个体
- 抽样和样本
- 随机样本与样本值

一. 总体和个体

引例： 有一箱螺丝钉，设螺丝钉的长度为随机变量$X \sim N(\mu,\sigma^2)$

其中 $\sigma = 0.01, \mu = E(X)$ 未知，由箱中随机的抽取16个螺丝钉，长度分别为随机变量 $X_1, X_2, \cdots, X_n (n = 16)$,

测量其长度为：

2.23，2.21，2.20，2.24，2.22，2.25，2.21，2.24，
2.25，2.23，2.35，2.21，2.24，2.23，2.25，2.22

试求：$\mu = E(X)$ 的估计值。

总体： 研究对象的某项数量指标取值的全体称为总体. 一个总体对应一个随机变量。

个体：将总体中的每个元素称为个体.

二. 抽样和样本

引例： 有一箱螺丝钉，设螺丝钉的长度为随机变量$X \sim N(\mu,\sigma^2)$

其中 $\sigma = 0.01, \mu = E(X)$ 未知，由箱中随机的抽取16个螺丝钉，长度分别为随机变量$X_1, X_2, \cdots, X_n (n = 16)$,

测量其长度为：

2.23，2.21，2.20，2.24，2.22，2.25，2.21，2.24，
2.25，2.23，2.35，2.21，2.24，2.23，2.25，2.22

试求：$\mu = E(X)$ 的估计值。

抽样： 为推断总体分布及各种特征，按一定规则从总体中随机抽取若干个体进行观察试验，以获得有关总体的信息，这一抽取过程称为 “抽样”。

样本： 被抽出的部分个体称为总体的一个样本。
样本中所包含的个体数目称为样本容量。

三. 随机样本与样本值

引例： 有一箱螺丝钉，设螺丝钉的长度为随机变量 $X \sim N(\mu, \sigma^2)$

其中 $\sigma = 0.01, \mu = E(X)$ 未知，由箱中随机的抽取16个螺丝钉，长度分别为随机变量 $X_1, X_2, \cdots, X_n (n = 16)$,

测量其长度为：

2.23，2.21，2.20，2.24，2.22，2.25，2.21，2.24，
2.25，2.23，2.35，2.21，2.24，2.23，2.25，2.22

试求：$\mu = E(X)$ 的估计值。

随机样本： 若 $X_1, X_2, \cdots X_n$ 相互独立且与总体 X 同分布，则称其为来自总体 X 的一个(简单随机)样本。

样本值： 样本的观察值 $x_1, x_2, \cdots x_n$ 称为样本值。

第六章 数理统计的基本概念

第一节 总体与随机样本

- 总体与个体
- 抽样和样本
- 随机样本与样本值

第六章 数理统计的基本概念

第一节 总体与随机样本

第二节 常用统计量

第三节 常用统计量的分布---抽样分布

数理统计研究解决的问题

$X \sim B(1,p)$

$X \sim B(n,p)$

$X \sim P(\lambda)$

$X \sim U(a,b)$

$X \sim E(\theta)$

$X \sim N(\mu,\sigma^2)$

设总体 $X \sim F(x,\theta)$, θ 是未知参数

概率论的任务： 用 $F(x,\theta)$ 计算 $P(x_1 < X \le x_2)$

统计学的任务： 对未知参数 θ 进行统计推断

统计推断的方法： 样本推断总体

具体做法：

$E(X), D(X)$

样本均值

样本方差

1) 抽样： $X_1, X_2, \cdots, X_n$ (独立且与 X 同分布)

$x_1, x_2, \cdots, x_n$

2) 构造统计量： $g(X_1, X_2, \cdots, X_n) \sim$ 抽样分布 (Ch6)

$g(x_1, x_2, \cdots, x_n)$

3) 统计推断： 对未知参数 θ 进行统计推断 $\begin{cases} \text{参数估计(Ch7)} \surd \\ \text{假设检验(Ch8)} \times \end{cases}$

第六章 数理统计的基本概念

第二节 常用统计量

- 两个重要统计量----样本均值，样本方差
- 其他常用统计量

1. 两个重要统计量

设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的一个样本,

$E(X)$

(1) 样本均值: $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ $\quad\quad$ $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{P} E(X)$

Ch5 大数定律

$D(X)=E[X-E(X)]^2$

(2) 样本方差: √ $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 无偏估计量

× $S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$ 有偏估计量

1. 两个重要统计量

$$D(X)=E[X-E(X)]^2$$

设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的一个样本，

(1) 样本均值： $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

(2) 样本方差： √ $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ × $S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$

无偏估计量　　有偏估计量

因为： $\sum_{i=1}^{n}(X_i-\bar{X})=\sum_{i=1}^{n}X_i-n\bar{X}=0$

所以，当 $\bar{X}$ 确定后，n个偏差 $X_1-\bar{X}, X_2-\bar{X}, \cdots, X_n-\bar{X}$

中只有 $n-1$个可以自由取值，而第 n 个则不能自由取值。

即自由度为 $n-1$。

$\sqrt{E(X^k)}$

2. 其他常用统计量

设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的一个样本，

(1) 样本均值： $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

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

(3) 样本标准差：$S = \sqrt{S^2} = \sqrt{\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$

(4) 样本 k 阶矩：$A_k = \dfrac{1}{n}\sum_{i=1}^{n} X_i^k \quad k = 1, 2, \ldots$

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{P} E(X^k)$$ Ch5 大数定律

小结

常用统计量

设 $X_1, X_2, \ldots, X_n$ 是来自总体 X 的一个样本，

样本均值：$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	$E(X)$	$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \xrightarrow{P} E(X)$ Ch5 大数定律
样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$	$D(X)$	
样本 k 阶矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$	$E(X^k)$	$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k \xrightarrow{P} E(X^k)$ Ch5 大数定律

第二章 一维随机变量及其分布

第四节 连续型随机变量及其分布

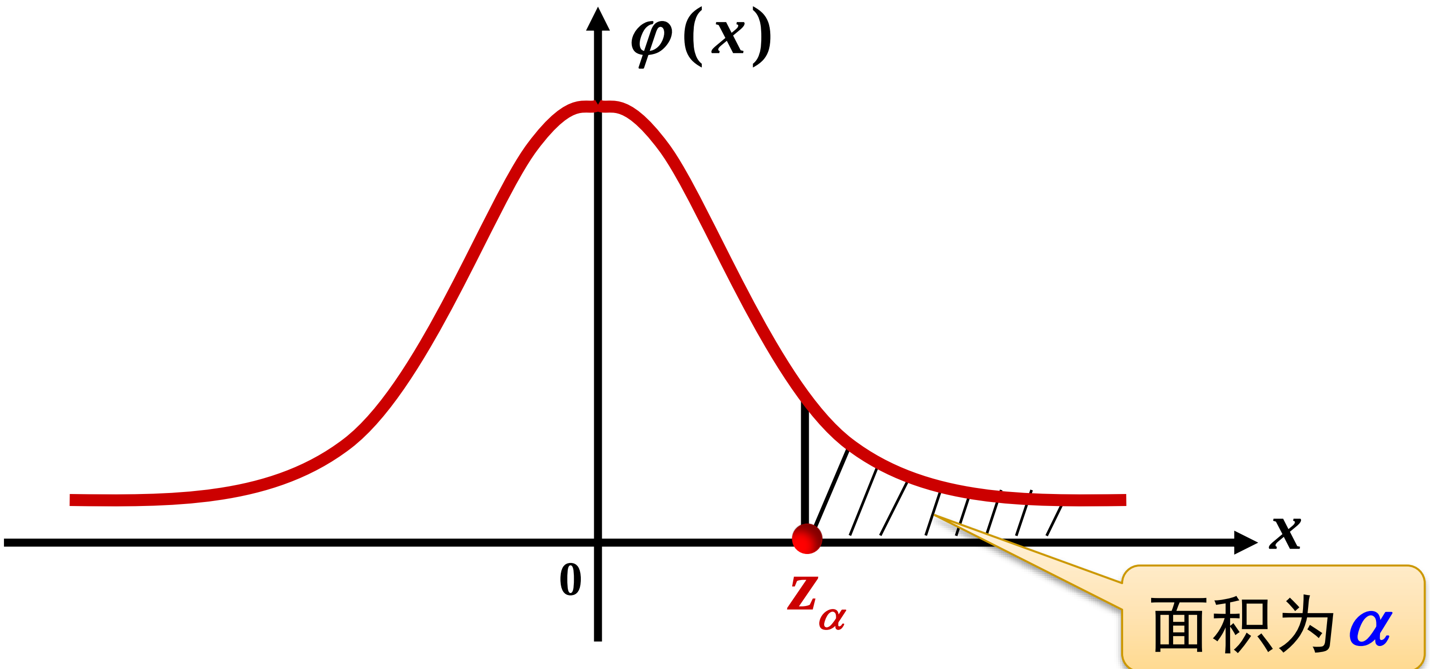
- 连续型随机变量的概率密度
- 几种常见的连续型随机变量的分布
  - 均匀分布
  - 指数分布
  - 正态分布
  - α分位点

4.α分位点

1) 定义 $X \sim N(0,1)$, 若 z_α 满足条件 $P(X > z_\alpha) = \alpha$,

$0 < \alpha < 1$, 则称点 z_α 为标准正态分布的上 α 分位点。

2) 图形

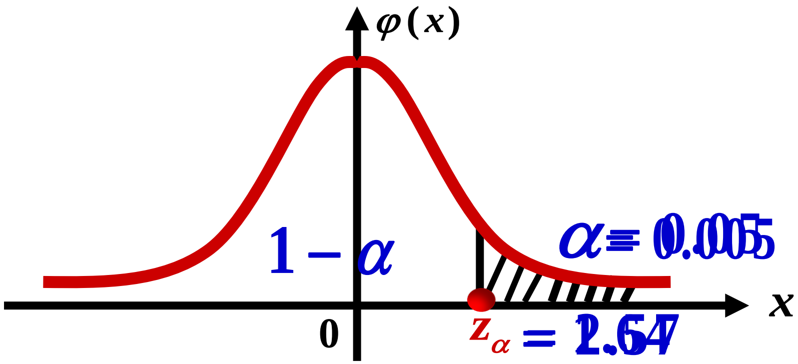


$$P(X > z_\alpha) = \int_{z_\alpha}^{+\infty} \varphi(x)\,\mathrm{d}x = \alpha$$

3) 分位点的计算：

▲ $\Phi(z_\alpha) = \int_{-\infty}^{z_\alpha} \varphi(x)\mathrm{d}x = 1 - \alpha$

（附表2上）



▲ 从正态分布表上如何求 z_α 的值:

$$\Phi(z_{0.05}) = 1 - 0.05 = 0.95 = \Phi(1.64) \qquad \therefore z_{0.05} = 1.64$$

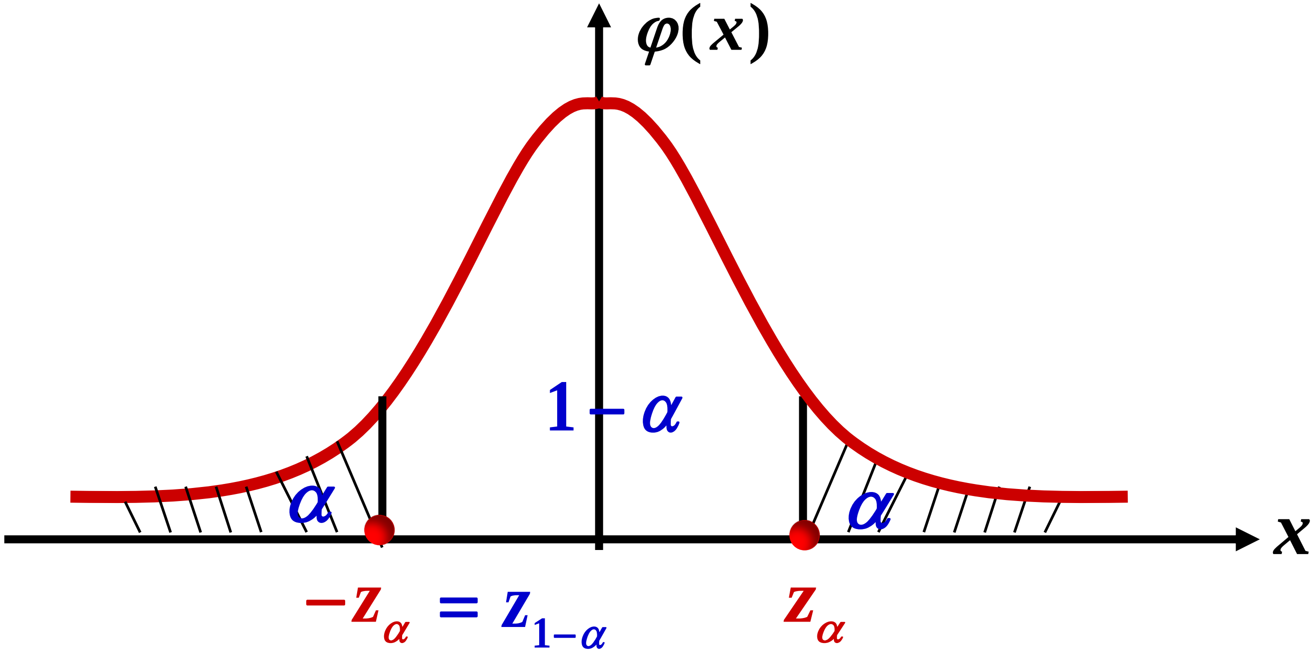
$$\Phi(z_{0.005}) = 1 - 0.005 = 0.995 = \Phi(2.57) \qquad \therefore z_{0.005} = 2.57$$

4) 分位点的性质：

性质1　α 越小，z_α 越大

性质2　$-z_\alpha = z_{1-\alpha}$

4) 分位点的性质： $-z_\alpha = z_{1-\alpha}$



注意: 在后续内容中还将介绍 $\chi^2(n)$ 分布，t分布的上 α 分位点的概念。

第六章 数理统计的基本概念

第一节 总体与随机样本

第二节 常用统计量

第三节 常用统计量的分布---抽样分布

第六章 数理统计的基本概念

$X \sim B(1, p)$

$X \sim B(n, p)$

$X \sim P(\lambda)$

$X \sim U(a,b)$

$X \sim E(\theta)$

$X \sim N(\mu,\sigma^2)$

第三节 常用统计量的分布

- 两个重要分布 χ^2分布 t 分布
- 正态总体样本均值与样本方差的分布

$X \sim N(\mu,\sigma^2)$ $X_1, X_2, \ldots, X_n$ 是一个样本

样本均值：$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ ~ 概率分布?

样本方差：$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ ~ 概率分布?

一. 两个重要的分布

1. χ^2 分布

$X_i \sim N(0,1)$ $i=1,2,\cdots,n$ 独立

定义 设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N(0,1)$的样本,

则称统计量: $\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$

为服从自由度为 n 的 χ^2 分布. 记为: $\chi^2 \sim \chi^2(n)$

统计量 概率密度

注: ▲ 自由度 n 是指 χ^2中所包含独立变量的个数

▲ $\chi^2 \sim \chi^2(1)$ $\chi^2 = X_1^2$ $X_1 \sim N(0,1)$

$$X \sim N(\mu,\sigma^2) \qquad f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$X_1, X_2, \ldots, X_n$ 独立且都 $\sim N(0,1)$ $\quad \chi^2 = X_1^{\,2} + X_2^{\,2} + \cdots + X_n^{\,2} \sim \chi^2(n)$

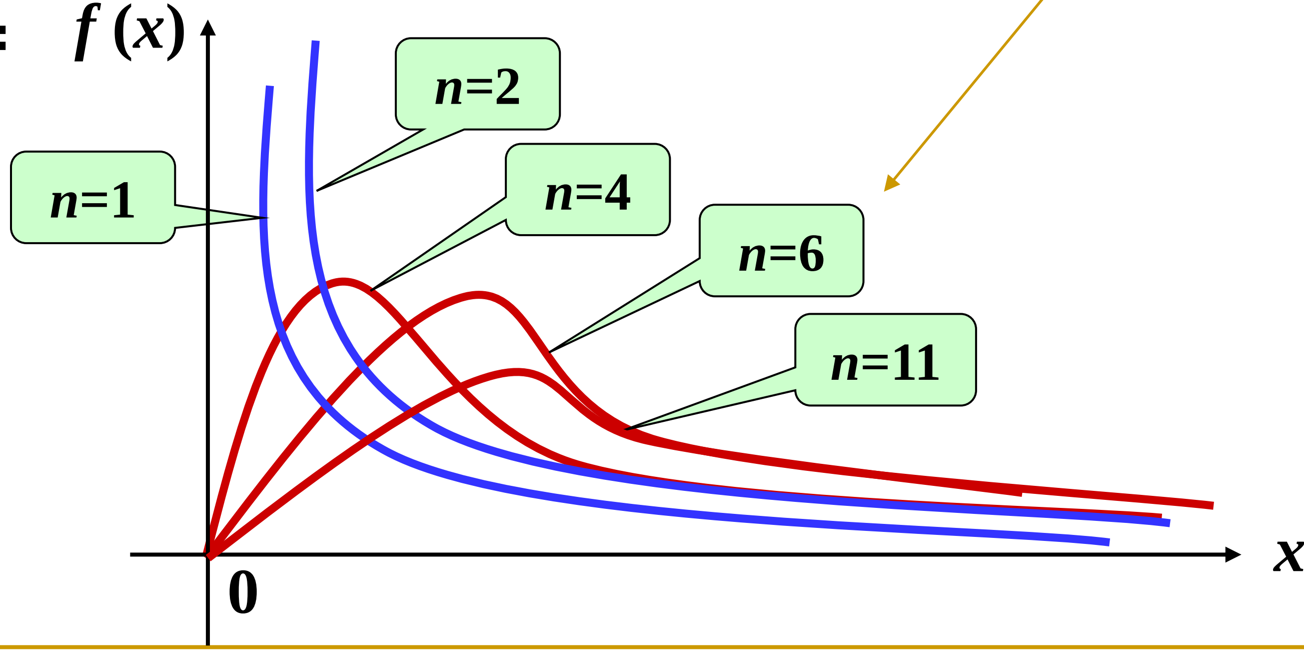
概率密度

注：▲ χ^2 分布的密度函数为：

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

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} \mathrm{d}t, \quad x > 0$$

图形：

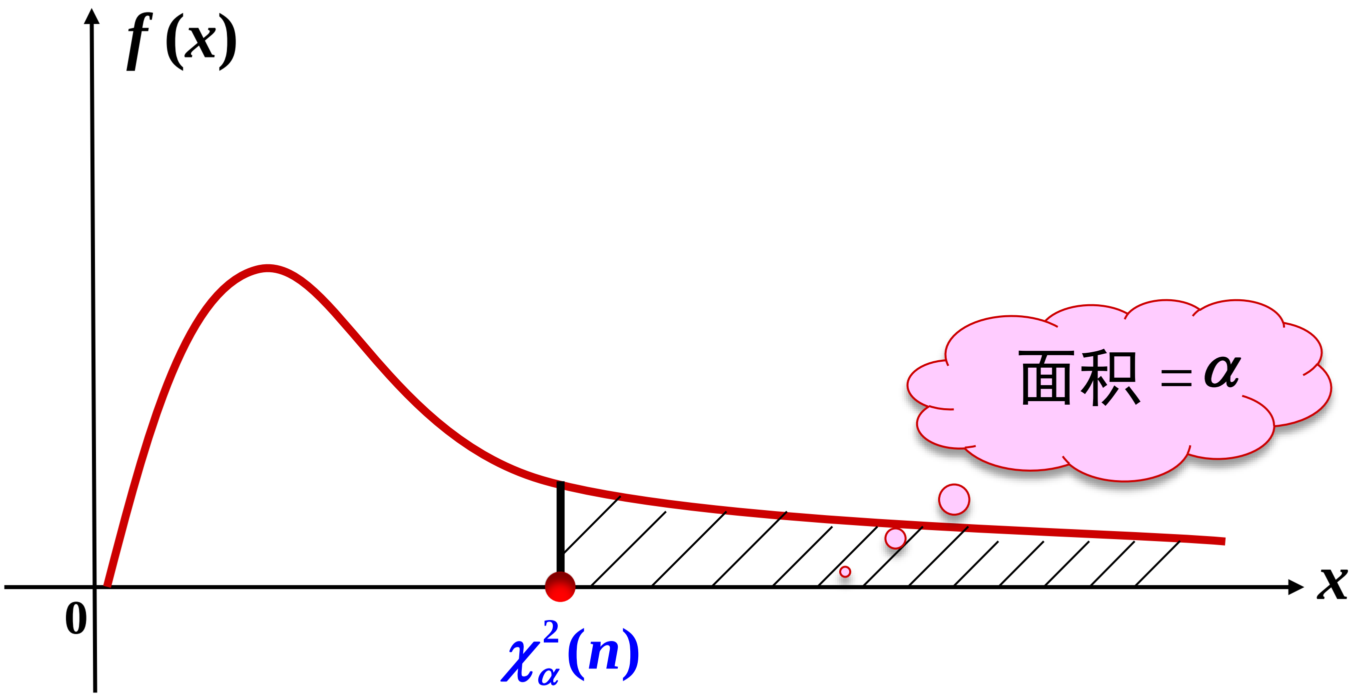


$\chi^2 \sim \chi^2(n)$

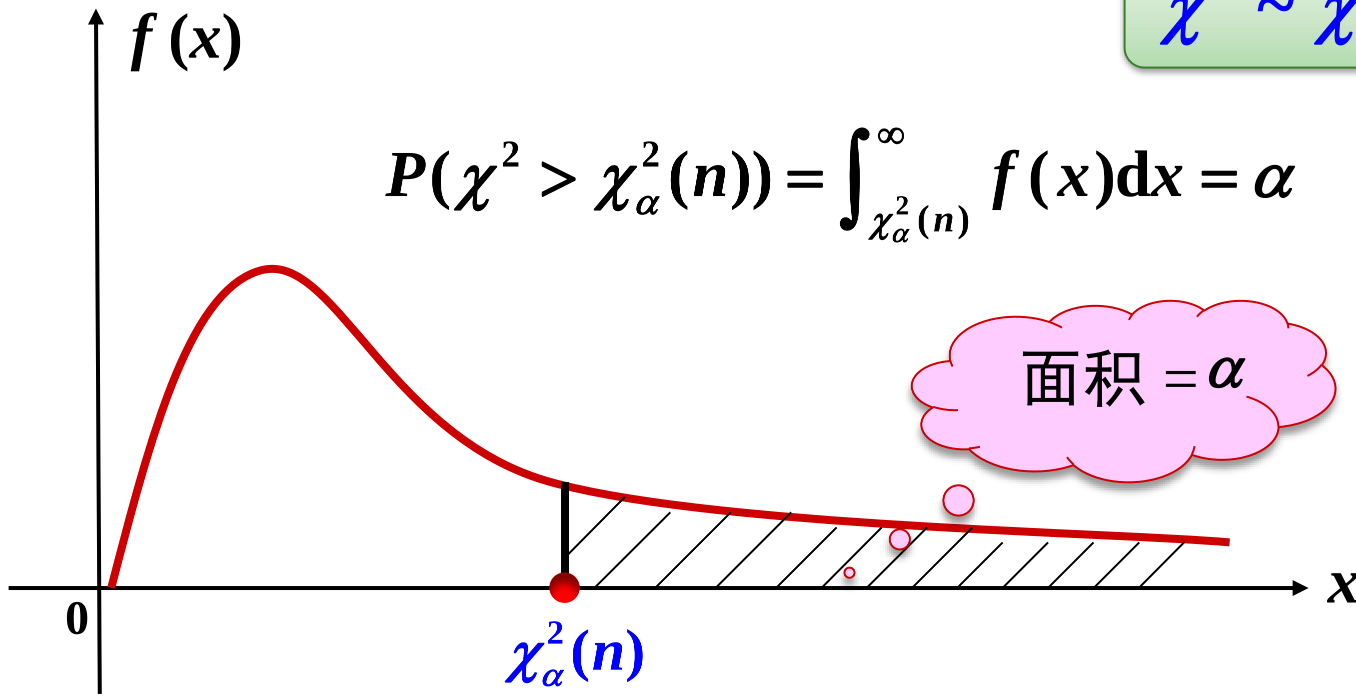
注：▲ χ^2分布的上 α 分位点：

对于给定的 α $(0<\alpha<1)$，称满足：

$P(\chi^2 > \chi^2_\alpha(n)) = \int_{\chi^2_\alpha(n)}^{\infty} f(x)\mathrm{d}x = \alpha$ 的点 $\chi^2_\alpha(n)$ 为 χ^2 分布的上α分位点。其图形如下：



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



注： χ^2, $\chi^2(n)$, $\chi^2_\alpha(n)$ 区别：

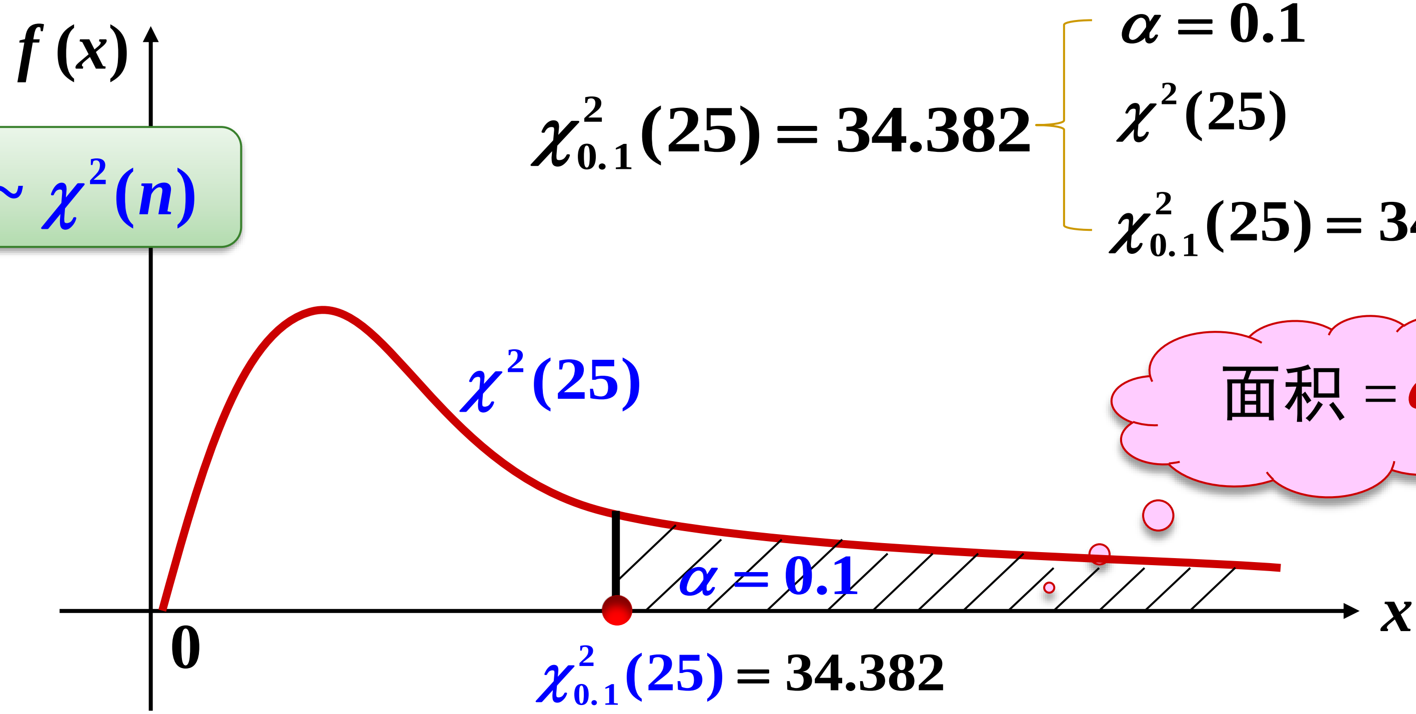
$\chi^2 = X_1{}^2 + X_2{}^2 + \cdots + X_n{}^2$ 是统计量；

$\chi^2(n)$ 是概率分布，表示概率密度是 $f(x)$；

$\chi^2_\alpha(n)$ 是分位点，是 x 轴上满足上述条件的一个实数.

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

$$\chi^2_{0.1}(25) = 34.382 \begin{cases} \alpha = 0.1 \\ \chi^2(25) \\ \chi^2_{0.1}(25) = 34.382 \end{cases}$$

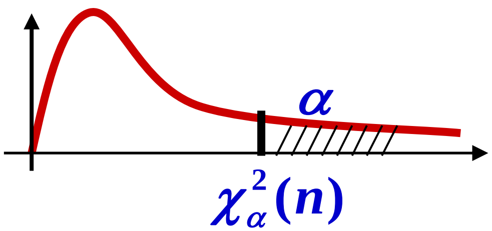


面积 $= \alpha$

对于不同的α与 n，$\chi^2_\alpha(n)$ 有表可查(见教材的附表5)

当 $n \leq 40$ 时可直接查表，例如：$\chi^2_{0.1}(25) = 34.382$

附表5 χ^2分布表---P386

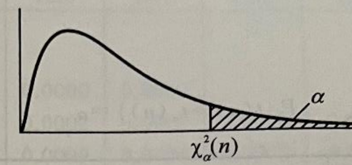


$$\chi^2_{0.1}(25) = 34.382$$

$$\chi^2_{0.95}(40) = 26.509$$

当$n > 40$时可用近似公式：

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

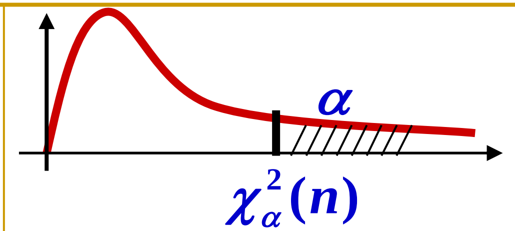


$P\{\chi^2(n) > \chi^2_\alpha(n)\} = \alpha$

n \ α	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	0.000	0.000	0.001	0.004	0.016	2.706	3.843	5.025	6.637	7.882
2	0.010	0.020	0.051	0.103	0.211	4.605	5.992	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.344	12.837
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.085	16.748
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.440	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.012	18.474	20.276
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.534	20.090	21.954
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.022	21.665	23.587
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.724	26.755
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.041	19.812	22.362	24.735	27.687	29.817
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.600	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.577	32.799
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.407	7.564	8.682	10.085	24.769	27.587	30.190	33.408	35.716
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.843	7.632	8.906	10.117	11.651	27.203	30.143	32.852	36.190	38.580
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.033	8.897	10.283	11.591	13.240	29.615	32.670	35.478	38.930	41.399
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.195	11.688	13.090	14.848	32.007	35.172	38.075	41.637	44.179
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.558
25	10.519	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.925
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.807	12.878	14.573	16.151	18.114	36.741	40.113	43.194	46.962	49.642
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.120	14.256	16.147	17.708	19.768	39.087	42.557	45.722	49.586	52.333
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
31	14.457	15.655	17.538	19.280	21.433	41.422	44.985	48.231	52.190	55.000
32	15.134	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486	56.328
33	15.814	17.073	19.046	20.866	23.110	43.745	47.400	50.724	54.774	57.646
34	16.501	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061	58.964
35	17.191	18.508	20.569	22.465	24.796	46.059	49.802	53.203	57.340	60.272
36	17.887	19.233	21.336	23.269	25.643	47.212	50.998	54.437	58.619	61.581
37	18.584	19.960	22.105	24.075	26.492	48.363	52.192	55.667	59.891	62.880
38	19.289	20.691	22.878	24.884	27.343	49.513	53.384	56.896	61.162	64.181
39	19.994	21.425	23.654	25.695	28.196	50.660	54.572	58.119	62.426	65.473
40	20.706	22.164	24.433	26.509	29.050	51.805	55.758	59.342	63.691	66.766

小结

常用统计量及抽样分布

<table>
<tr><td>χ^2分布</td><td>$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立
★$\chi^2=\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$</td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

第六章 数理统计的基本概念

第三节 常用统计量的分布

- 两个重要分布 χ^2分布 t 分布
- 正态总体样本均值与样本方差的分布

2. t 分布

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

则称统计量：$t = \dfrac{X}{\sqrt{Y/n}}$

为服从自由度为 n 的 t 分布．记为：$t \sim t(n)$

概率密度

注: ➢ t 分布是英国统计学家哥塞特(Gosset)首先发现的，并以学生(*student*)的笔名在英国的《*Biometrike*》杂志上发表的一篇论文中提出了他的研究结果，故 t 分布也称为学生分布。

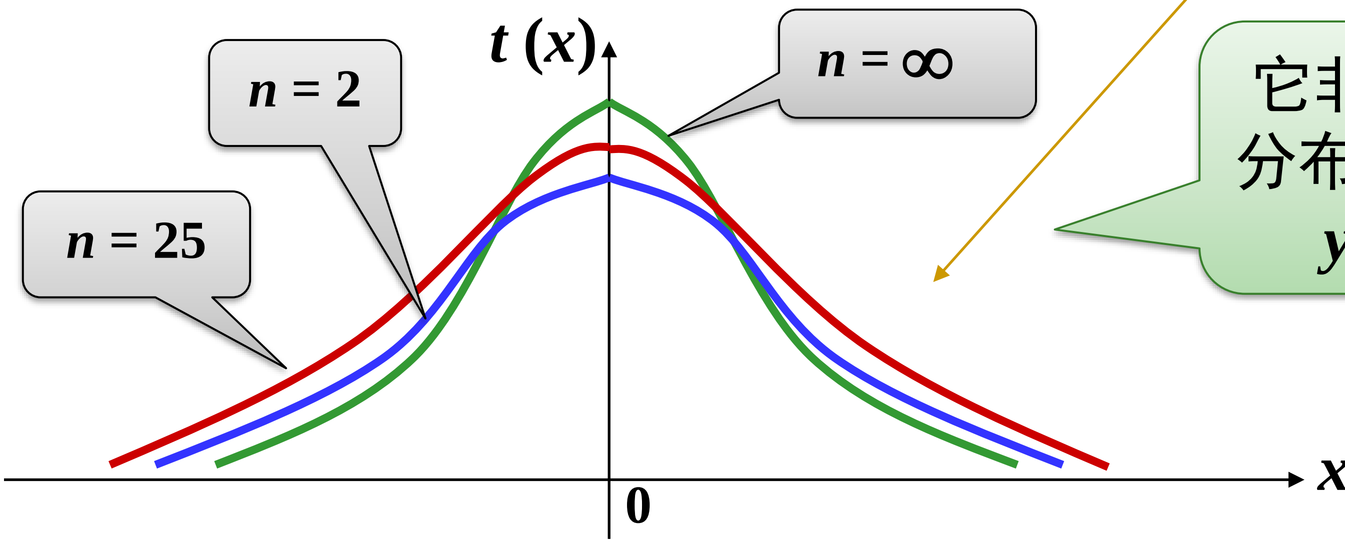
$$t = \frac{X}{\sqrt{Y/n}} \sim t(n)$$

概率密度

➢ t 分布的概率密度函数为：

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

其图形如下：



它非常象正态分布图形，关于 y 轴对称

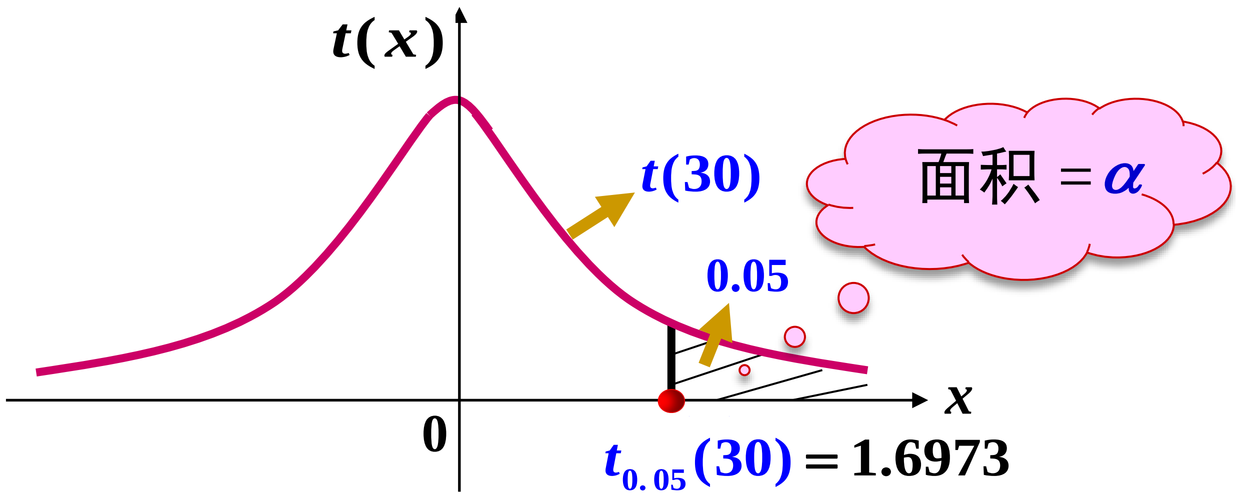
➢可以证明：$\lim\limits_{n\to\infty} t(x;n) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \varphi(x)$

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

➢ t 分布的上 α 分位点：对于给定的 α，$(0<\alpha<1)$

称满足条件：$P(t > t_\alpha(n)) = \int_{t_\alpha(n)}^{\infty} t(x)\,\mathrm{d}x = \alpha$

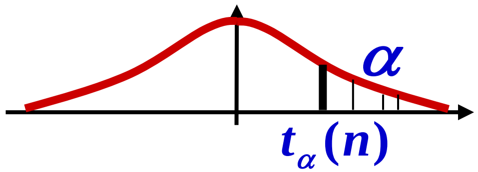
的点 $t_\alpha(n)$ 为 t 分布的上 α 分位点。



对于不同的 α 与 n，$t_\alpha(n)$ 有表可查(见教材的附表4)

例如： $t_{0.05}(30) = 1.6973$

附表4　t分布表---P385



$$t_{0.05}(30) = 1.6973$$

$$t_{0.01}(35) = 2.4377$$

当$n > 45$时可用近似公式：

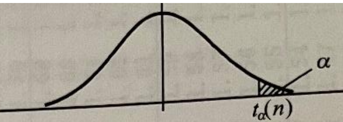
$$t_\alpha(n) \approx z_\alpha$$

$$t_{0.05}(50) \approx z_{0.05}$$

$$\Phi(z_{0.05}) = 0.95$$

$$z_{0.05} = 1.645$$

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



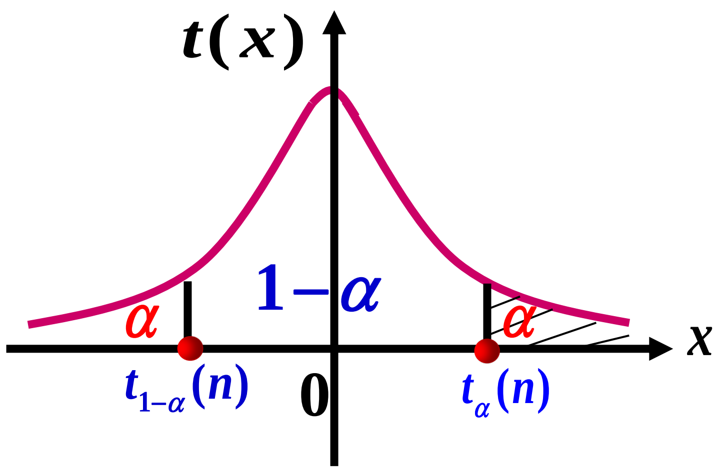
n \ α	0.20	0.15	0.10	0.05	0.025	0.01	0.005
1	1.376	1.963	3.0777	6.3138	12.7062	31.8207	63.6574
2	1.061	1.386	1.8856	2.9200	4.3027	6.9646	9.9248
3	0.978	1.250	1.6377	2.3534	3.1824	4.5407	5.8409
4	0.941	1.190	1.5332	2.1318	2.7764	3.7469	4.6041
5	0.920	1.156	1.4759	2.0150	2.5706	3.3649	4.0322
6	0.906	1.134	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.896	1.119	1.4149	1.8946	2.3646	2.9980	3.4995
8	0.889	1.108	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.883	1.100	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.879	1.093	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.876	1.088	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.873	1.083	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.870	1.079	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.868	1.076	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.866	1.074	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.865	1.071	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.863	1.069	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.862	1.067	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.861	1.066	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.860	1.064	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.859	1.063	1.3232	1.7207	2.0796	2.5177	2.8314
22	0.858	1.061	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.858	1.060	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.857	1.059	1.3178	1.7109	2.0639	2.4922	2.7969
25	0.856	1.058	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.856	1.058	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.855	1.057	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.855	1.056	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.854	1.055	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.854	1.055	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.8535	1.0541	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.8531	1.0536	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.8527	1.0531	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.8524	1.0526	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.8521	1.0521	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.8518	1.0516	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.8515	1.0512	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.8512	1.0508	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.8510	1.0504	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.8507	1.0501	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.8505	1.0498	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.8503	1.0494	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.8501	1.0491	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.8499	1.0488	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.8497	1.0485	1.3006	1.6794	2.0141	2.4121	2.6896

t 分布分位点的性质:

$$P(t > t_\alpha(n)) = \int_{t_\alpha(n)}^{\infty} t(x)\mathrm{d}x = \alpha$$

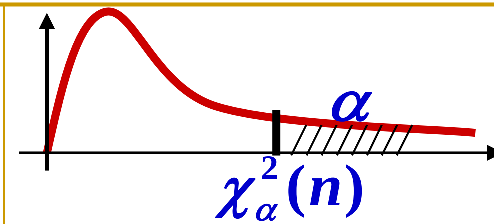
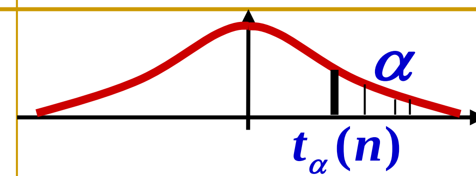
➤由上 α 分位点定义及 $t(x)$ 对称性得:

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



小结

常用统计量及抽样分布

χ^2分布	$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立 ★$\chi^2=\sum\limits_{i=1}^{n}X_i^2 \sim \chi^2(n)$	
t 分布	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立 ★$t=\dfrac{X}{\sqrt{Y/n}} \sim t(n)$	 $t_\alpha(n)=-t_{1-\alpha}(n)$

练习

1. 设 X_1, X_2, X_3, X_4, X_5 是来自正态总体 $N(0,1)$的样本，

统计量 $X_1^2+X_2^2+X_3^2+X_4^2+X_5^2$ 服从 (D) 分布？

(A) $t(4)$ (C) $\chi^2(4)$

(B) $t(5)$ (D) $\chi^2(5)$

$X_i \sim N(0,1)$ 独立 $i=1,2,\cdots,n$

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

2. 设 X_1, X_2, X_3, X_4, X_5 是来自正态总体 $N(0,1)$的样本，

服从 $\chi^2(3)$ 分布的统计量是 (C) (D)

(A) $X_1+X_2+X_3$ (C) $X_1^2+X_2^2+X_3^2$

(B) $X_1^2-X_2^2+X_3^2$ (D) $X_3^2+X_4^2+X_5^2$

练习

3. 设 $X \sim N(0,1)$，$Y \sim \chi^2(4)$，且 X 与 Y 相互独立，

统计量 $\dfrac{X}{\sqrt{Y/4}}$ 服从 ___(A)___ 分布？

(A) $t(4)$　　(C) $\chi^2(4)$

(B) $t(5)$　　(D) $\chi^2(5)$

$X \sim N(0,1)$,

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

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

4. 设 $X \sim N(0,1)$，$Y \sim \chi^2(3)$，且 X 与 Y 相互独立，

服从 $t(3)$ 分布的统计量是 ___(B)___

(A) $\dfrac{X}{\sqrt{Y/4}}$　　(C) $\dfrac{Y}{\sqrt{X/4}}$

(B) $\dfrac{X}{\sqrt{Y/3}}$　　(D) $\dfrac{Y}{\sqrt{X/3}}$

3. F分布

定义 设 $X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, X 与Y 相互独立,

则称统计量: $F = \dfrac{X/n_1}{Y/n_2}$

为服从自由度 n_1 及 n_2 的 F 分布, 记作: $F \sim F(n_1, n_2)$

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

证明: $\because$ $X \sim t(n)$，所以由 t 分布的定义，

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

$U \sim N(0,1) \rightarrow U^2 \sim \chi^2(1)$

$V \sim \chi^2(n)$　且 U, V 独立

$$X^2 = \frac{U^2}{\frac{V}{n}} = \frac{\frac{U^2}{1}}{\frac{V}{n}} \sim F(1,n)$$ (由 F 分布定义)

$X \sim \chi^2(n_1)$,

$Y \sim \chi^2(n_2)$, 独立

$$F = \frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$$

第六章 数理统计的基本概念

第三节 常用统计量的分布

- 两个重要分布 χ^2分布 t 分布
- 正态总体样本均值与样本方差的分布

$X \sim N(\mu, \sigma^2)$ $X_1, X_2, \ldots, X_n$ 是一个样本

样本均值： $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim$ 概率分布?

样本方差： $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \sim$ 概率分布?

二. 正态分布的样本均值与样本方差的分布

$X_i \sim N(\mu,\sigma^2)$

定理1 设 $X_1, X_2,\ldots, X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的样本，

$\bar{X}, S^2$ 是其样本均值和样本方差

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

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

解：$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \sim N(\underset{\mu}{E\bar{X}}, \underset{\frac{\sigma^2}{n}}{D\bar{X}})$

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

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

$$\boxed{k_1X_1 + k_2X_2 + \cdots + k_nX_n \sim N(k_1\mu_1 + k_2\mu_2 + \cdots k_n\mu_n, k_1^2\sigma_1^2 + k_2^2\sigma_2^2 + \cdots + k_n^2\sigma_n^2)}$$

二. 正态分布的样本均值与样本方差的分布

$X_i \sim N(\mu,\sigma^2)$

定理1 设 $X_1, X_2, \ldots, X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的样本，

$\bar{X}, S^2$ 是其样本均值和样本方差

$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

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

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

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

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

$$\rightarrow \frac{(n-1)S^2}{\sigma^2}=\sum_{i=1}^{n}(\underbrace{\frac{X_i-\overset{\mu}{\bar{X}}}{\sigma}}_{N(0,1)})^2 \sim \chi^2(n-1)$$

$X_1-\bar{X}, X_2-\bar{X}, \cdots, X_n-\bar{X}$ 自由度为 $n-1$。

二. 正态分布的样本均值与样本方差的分布

定理1 设 $X_1, X_2, \dots, X_n$ 是取自正态总体 $N(\mu, \sigma^2)$ 的样本，

$\bar{X}, S^2$ 是其样本均值和样本方差

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

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

(3) $\bar{X}$ 和 S^2 相互独立

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

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

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

二. 正态分布的样本均值与样本方差的分布

定理2 设 $X_1, X_2, \dots, X_n$ 是取自正态总体 $N(\mu,\sigma^2)$ 的样本，$\bar{X}$ 和 S^2 分别为样本均值和样本方差，

$$\boxed{t=\frac{U}{\sqrt{V/n-1}}\sim t(n-1)}$$

则有：

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

$U\sim N(0,1)$，$V\sim\chi^2(n-1)$，独立

证明： U与V相互独立，由t分布的定义得：

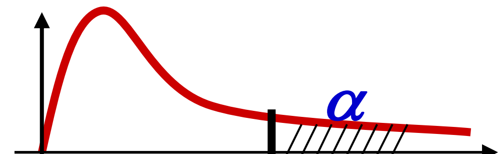
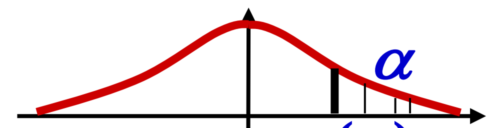
$$t=\frac{U}{\sqrt{V/n-1}}=\frac{\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S^2}{(n-1)\sigma^2}}}=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$

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

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

小结

常用统计量及抽样分布

χ^2 分布	$X_i \sim N(0,1)\ \ i=1,2,\cdots,n$ 独立 ★ $\chi^2=\sum\limits_{i=1}^{n} X_i^2 \sim \chi^2(n)$	
t 分布	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立 ★ $t=\dfrac{X}{\sqrt{Y/n}} \sim t(n)$	
$X \sim N(\mu,\sigma^2)$ $X_1,X_2,\cdots,X_n$ $\overline{X},S^2$ $\overline{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$	★*Th1* $\overline{X} \sim N(\mu,\sigma^2/n)$, 独立 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$	*Th2* ★ $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

小结

常用统计量及抽样分布

$X \sim N(\mu,\sigma^2)$

$X_1, X_2, \cdots, X_n$

	统计量	概率分布	
χ^2统计量	$\chi^2 = \sum_{i=1}^{n} X_i^2$	$\sim \chi^2(n)$	$X_i \sim N(0,1)\ i=1,2,\cdots,n$ 独立
t 统计量	$t = \dfrac{X}{\sqrt{Y/n}}$	$\sim t(n)$	$X \sim N(0,1),\ Y \sim \chi^2(n)$, 独立
样本均值	$\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$	$\sim N(\mu, \dfrac{\sigma^2}{n})$	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
样本方差	$\dfrac{(n-1)S^2}{\sigma^2}$	$\sim \chi^2(n-1)$	$S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$
	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\sim t(n-1)$	

练习

1. 设 X_1, X_2, X_3, X_4, X_5 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，

样本均值 $\bar{X}=\frac{1}{5}\sum_{i=1}^{5}X_i$ 服从 (D) 分布？

(A) $t(5)$

(B) $\chi^2(5)$

(C) $N(\mu,\sigma^2/4)$

(D) $N(\mu,\sigma^2/5)$

$$\boxed{\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\sim N(\mu,\frac{\sigma^2}{n})}$$

2. 设 X_1, X_2, X_3, X_4, X_5 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，

服从 $N(\mu,\sigma^2/5)$ 分布的统计量是 (C)

(A) $X_1+X_2+X_3$

(B) $S^2=\frac{1}{4}\sum_{i=1}^{n}(X_i-\bar{X})^2$

(C) $\bar{X}=\frac{1}{5}\sum_{i=1}^{n}X_i$

(D) $X_3^2+X_4^2+X_5^2$

练习

3. 设 X_1, X_2, X_3, X_4, X_5 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，

统计量 $\dfrac{4S^2}{\sigma^2}$ 服从 (A) 分布？

(A) $\chi^2(4)$

(B) $\chi^2(5)$

(C) $t(4)$

(D) $t(5)$

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

4. 设 X_1, X_2, X_3, X_4是来自正态总体 $N(\mu,\sigma^2)$ 的样本，

服从 $\chi^2(3)$ 分布的统计量是 (B)

(A) $\dfrac{X}{\sqrt{Y/4}}$

(B) $3S^2/\sigma^2$

(C) $\bar{X} = \dfrac{1}{4}\sum_{i=1}^{n} X_i$

(D) $X_1^2 + X_2^2 + X_3^2$

练习

5. 设 X_1, X_2, X_3, X_4, X_5 是来自正态总体 $N(\mu,\sigma^2)$ 的样本，

统计量 $\dfrac{\bar{X}-\mu}{S/\sqrt{5}}$ 服从 (C) 分布？

$(A)\ \chi^2(4)$ $(C)\ t(4)$

$(B)\ \chi^2(5)$ $(D)\ t(5)$

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

6. 设 X_1, X_2, X_3, X_4, X_5是来自正态总体 $N(\mu,\sigma^2)$ 的样本，

统计量 $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{5}}$ 服从 (D) 分布?

$(A)\ \chi^2(4)$ $(C)\ t(4)$

$(B)\ \chi^2(5)$ $(D)\ N(0,1)$

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

第六章 数理统计的基本概念

作 业

授课内容	习题六
6.1 随机样本	1, 3
6.3 抽样分布	4（1）（2）， 5（1）（2），9（1）