

第七章 分析化学中的数据处理

7.3 少量数据的统计处理(重点)

7.3.1 t分布曲线

7.3.2 平均值的置信区间

7.3.3 显著性检验

7.3.4 异常值的取舍

7.4 误差的传递

7.4.1 系统误差的传递

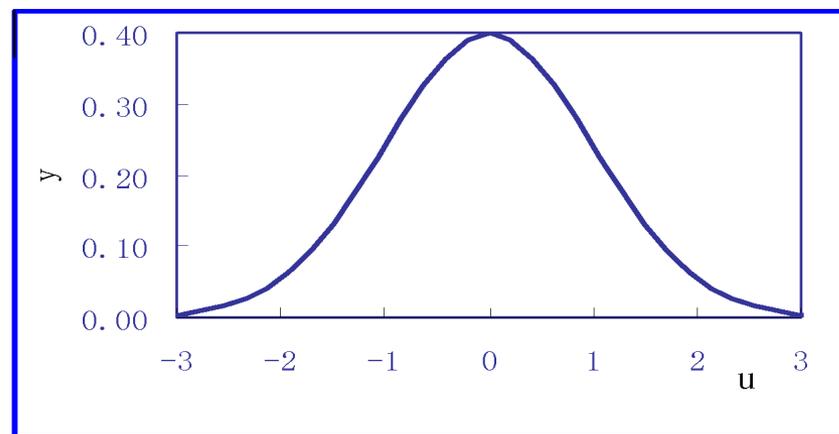
7.4.2 随机误差的传递

7.3 少量数据的统计处理

正态分布是无限次测量数据的分布规律，而在实际工作中，只能对随即抽得的样本进行有限次的测量。如何以统计的方法处理有限次测量数据，使其能合理地推断总体的特性。这是本次课要讨论的问题。

无限次测量，得到

$$\mu \quad \sigma \quad u = \frac{x - \mu}{\sigma}$$
$$y = \Phi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$$



u 分布曲线

7.3.1 t分布曲线

当测量数据不多时，总体标准偏差 σ 是不知道的，只好用样本标准偏差 s 来估计测量数据的分散情况。用 s 代替 σ ，必然引起正态分布的偏离，这时可用 t 分布来处理。

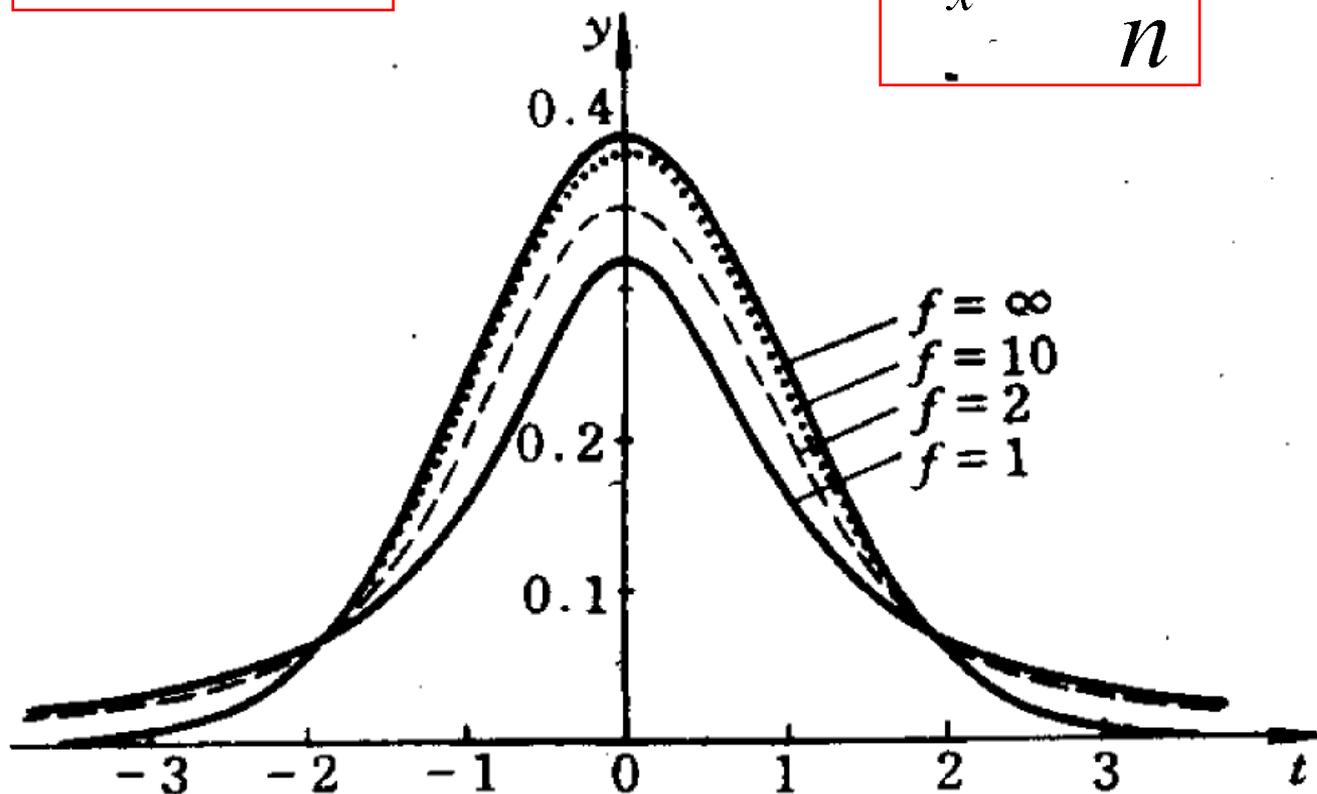
有限次测量，用 s 代替 σ

$$t = \frac{x - \mu}{s_{\bar{x}}} \rightarrow x = \mu + s_{\bar{x}} t \quad \text{代入}$$

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

$$f = n - 1$$

$$S_{\bar{x}} = \frac{S}{n}$$



t 分布曲线

t 分布曲线与正态分布曲线相似，曲线下面一定区间的面积表示随即误差出现的概率。只是 t 分布曲线随自由度 f 而改变。当 f 趋近无穷时， t 分布就趋近正态分布。

置信度P

P表示某一t值时，测定值落在 $(\mu \pm ts)$ 范围内的概率。落在此范围之外的概率为 $(1-P)$ ，称为显著性水准，用 α 表示。t值一般与置信度及自由度有关，一般表示为 $t_{\alpha, f}$ 。

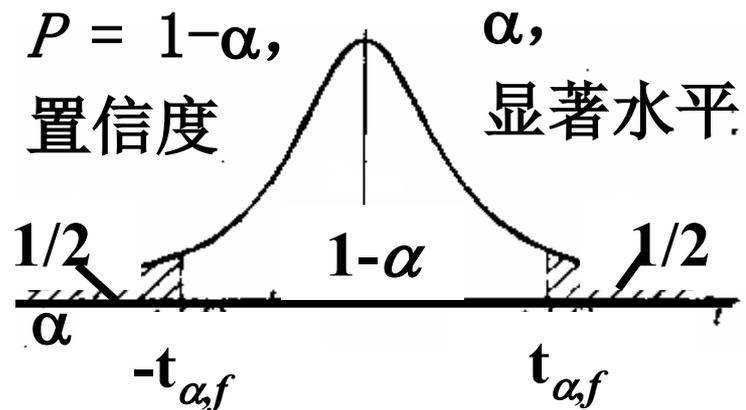
例如： $t_{0.05, 10}$ ， $t_{0.01, 5}$

$t_{0.05, 10}$ 表示置信度为95%，自由度为10时的t值；

$t_{0.01, 5}$ 表示置信度为99%，自由度为5时的t值。

t分布值表

自由度 $f=(n-1)$	显著水平 α			
	0.50	0.10	0.05	0.01
1	1.00	6.31	12.71	63.66
2	0.82	2.92	4.30	9.93
3	0.76	2.35	3.18	5.84
4	0.74	2.13	2.78	4.60
5	0.73	2.02	2.57	4.03
6	0.72	1.94	2.45	3.71
7	0.71	1.90	2.37	3.50
8	0.71	1.86	2.31	3.36
9	0.70	1.83	2.26	3.25
10	0.70	1.81	2.23	3.17
20	0.69	1.73	2.09	2.85
∞	0.67	1.65	1.96	2.58



6次测量，随机误差落在 ± 2.57 范围内的概率为95%。

无限次测量，随机误差落在 $\pm 1.96\sigma$ 范围内的概率为95%。

t分布值表

单位为

$$S_{\bar{x}}$$

单位为

 σ
还原为
u 分布

自由度 $f=(n-1)$	显著水平 α			
	0.50	0.10	0.05	0.01
1	1.00	6.31	12.71	63.66
2	0.82	2.92	4.30	9.93
3	0.76	2.35	3.18	5.84
4	0.74	2.13	2.78	4.60
5	0.73	2.02	2.57	4.03
6	0.72	1.94	2.45	3.71
7	0.71	1.90	2.37	3.50
8	0.71	1.86	2.31	3.36
9	0.70	1.83	2.26	3.25
10	0.70	1.81	2.23	3.17
20	0.69	1.73	2.09	2.85
∞	0.67	1.65	1.96	2.58

7.3.2 平均值的置信区间

——对 μ 的区间的估计

对一样品分析，报告出： \bar{x}, s, n 例如 $w_B = \bar{x} \pm s(n)$

无限次测量 $n \rightarrow \infty, \bar{x} \rightarrow \mu$

问题：对有限次测量如何通过 \bar{x} 估计 μ

在 \bar{x} 的某个范围内包含 μ 的概率有多大？

(..... \bar{x})

这个问题涉及两个方面：

- 1、区间界限，多大区间？ 置信区间问题
- 2、概率多大？ 置信度问题

例: $\bar{x} = 80.00$

μ 包含在 区间

$$80.00 \pm 0.15$$

置信度
↓
概率相对大

置信区间

$$80.00 \pm 0.05$$

概率相对小

$$80.00 \pm \infty$$

概率为100%，无意义

定义: 表示在一定置信度下, 以平均值 \bar{x} 为中心, 包括总体平均值的范围。

对于少量测量数据, 按照t的定义可得出:

$$\mu = \bar{x} \pm ts_{\bar{x}} = \bar{x} \pm \frac{ts}{\sqrt{n}}$$

例2. 分析铁矿中的铁的质量分数，得到如下数据：

37.45, 37.20, 37.50, 37.30, 37.25 (%)。

求：置信度分别为95%和99%的置信区间。

解： $\mu = \bar{x} \pm tS_{\bar{x}} = \bar{x} \pm \frac{tS}{\sqrt{n}}$ 需要计算那些量？

$$\begin{aligned}\bar{x} &= \frac{37.45 + 37.20 + 37.50 + 37.30 + 37.25}{5} \% \\ &= 37.34\%\end{aligned}$$

$$\begin{aligned} s &= \sqrt{\frac{\sum d_i^2}{n-1}} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \\ &= \sqrt{\frac{(0.11)^2 + (0.14)^2 + (0.04)^2 + (0.16)^2 + (0.09)^2}{5-1}} \\ &= 0.13\% \end{aligned}$$

还有t怎么求？

置信度为95%，即 $1-\alpha = 0.95$ ， $\alpha = 0.05$ ，

查表 $t_{0.05, 4} = 2.78$

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/958044103010006123>