揚州大學
YANGZHOU UNIVERSITY

生物统计与试验设计

Biostatistics and Experimental Design

主讲：杨泽峰

扬州大学农学院

第二章 描述性统计数

➢2.1 变量与次数分布

➢2.2 次数分布表

➢2.3 次数分布图

➢2.4 集中趋势的统计数

➢2.5 离散趋势的统计数

2.5 离散趋势的统计数

变异数：描述变量离散特性的统计数。

常用的变异数主要有：

① 极差

② 方差

③ 标准差

④ 变异系数

2.5.1 方差和标准差

扬州大学 YANGZHOU UNIVERSITY

（1）总体的方差和标准差

方差（variance）和标准差（standard deviation）都是反应变量的变异程度的。

设有一具有 N 个个体 y_1，y_2，…，y_N，且平均数为 μ 的有限总体，其总体方差 σ^2 为：

$$\sigma^2=\frac{1}{N}(y_1-\mu)^2+\frac{1}{N}(y_2-\mu)^2+\cdots\frac{1}{N}(y_n-\mu)^2=\frac{1}{N}\sum_{i=1}^{n}(y_i-\mu)^2$$

总体标准差 σ 定义为：

$$\sigma=\sqrt{\frac{\sum(y-\mu)^2}{N}}$$

2.5.1 方差和标准差

（2）样本的方差和标准差

设有一具有 n 个观察值 y_1， y_2， …， y_n， 且平均数为 $\bar{y}$ 的样本，其对总体方差 σ^2 和总体标准差 σ 的相应估计值分别为**样本方差** s^2 和**样本标准差** s 分别为：

$$s^2=\frac{\sum(y-\bar{y})^2}{n-1}，\quad s=\sqrt{\frac{\sum(y-\bar{y})^2}{n-1}}$$

其中：$\sum(y-\bar{y})^2$ 是离均差的平方和，简称**平方和**（sum of squares)，记作 SS； $(n-1)$ 称为**自由度**（degree of freedom)，在统计上指的是独立观察值的数目，记作 df 或 ν 。

平方和在计算中可以用恒等式： $SS=\sum(y_i-\bar{y})^2=\sum y_i^2-\frac{\left(\sum y_i\right)^2}{n}$

2.5.1 方差和标准差

(2）样本的方差和标准差

【例2-8】测得某品种10个玉米果穗质量（g）分别为：170，160，172，162，165，175，171，168，160，151，试计算其标准差。

表 2-7　玉米果穗质量数据资料整理

编号	y_i	y_i^2	$y_i-\bar{y}$	$(y_i-\bar{y})^2$
1	170	28900	4.6	21.16
2	160	25600	-5.4	29.16
3	172	29584	6.6	43.56
4	162	26244	-3.4	11.56
5	165	27225	-0.4	0.16
6	175	30625	9.6	92.16
7	171	29241	5.6	31.36
8	168	28224	2.6	6.76
9	160	25600	-5.4	29.16
10	151	22801	-14.4	207.36
求和	1654	274044	0	472.4

样本平均数： $\bar{y}=\dfrac{\sum y}{n}=\dfrac{1654}{10}=165.4$

平方和： $SS=\sum(y_i-\bar{y})^2=472.4$

$$SS=\sum y_i^2-\frac{\left(\sum y_i\right)^2}{n}=274044-\frac{1654^2}{10}=472.4$$

标准差： $s=\sqrt{\dfrac{SS}{\nu}}=\sqrt{\dfrac{472.4}{10-1}}\approx 7.24$ （g）

该品种玉米果穗重量的标准差为 7.24g。

2.5.1 方差和标准差

(3) 标准差的特性

① 标准差的大小受每个观察值的影响，如观察值间变异大，求得的标准差也大，反之则小。

② 各观察值加上或减去一个常数，标准差不变。

③ 每个观察值乘以或除以一个常数 a （$a>0$），则所得的标准差是原来标准差的 a 倍或 $1/a$ 倍。

2.5.2 变异系数

标准差是变量的平均变异量。

变量的相对变异量叫做**变异系数**（coefficient of variation），记作 *CV*。 $CV = \frac{s}{\bar{y}} \times 100\%$

【例 2-9】某品种水稻在大田种植时，其每穗粒数的平均数为 45.0，标准差为 17.8；而在丰产田种植时，其每穗粒数的平均数为 65.0，标准差为 18.4。试问哪种水稻田种植的水稻每穗粒数的变异程度较大?

大田种植变异系数： $CV_1 = \frac{s_1}{\bar{y}_1} \times 100\% = \frac{17.8}{45.0} \times 100\% = 39.56\%$

丰产田种植变异系数： $CV_2 = \frac{s_2}{\bar{y}_2} \times 100\% = \frac{18.4}{65.0} \times 100\% = 28.31\%$

丰产田种植的水稻每穗粒数变异系数较小，整齐度由于大田种植。