试验统计手册

刘醒亚 编译

SYTISC

机械工业出版社

试 验 统 计 手 册

刘 醒 亚 编译

机 械 工 业 出 版 社

内 容 简 介

本书内容分三部分：一是试验统计公式及其推导；二是试验统计的英文词汇及常用的符号；三是重要的统计用表。

本书可供从事工业、农业、商业、其他各部门、统计局的统计工作人员及高等院校有关专业师生参考。

试验统计手册

刘 醒 亚 编译

*

责任编辑：贾 欣　　责任校对：刘思碚

封面设计：方 芬　　版式设计：冉晓华

*

机械工业出版社出版（北京阜成門外百万庄南街一号）

（北京市书刊出版业营业許可証出字第 117 号）

机械工业出版社印刷厂印刷

新华书店北京发行所发行·新华书店经售

*

开本 $787\times1092^{1}/_{32}$·印张$4^{5}/_{8}$·字数101千字

1990年2月北京第一版·1990年2月北京第一次印刷

印数 001—990·定价：4.20元

*

社科新书目：237-022

ISBN 7-111-01880-X/F·255

编译者的话

试验统计是应用统计方法来设计试验与分析试验数据的一门科学，也是考验假设是否正确的一种科学工具。它可以应用于工业、农业以及社会经济各个领域。正因为如此，试验统计在我国的科研和国民经济的许多领域中应用越来越广泛，日益受到各界人士的重视，国外许多科技专业书刊和文献报告中也大量涉及这方面的科学内容。为了满足社会急需，编译了这本小册子。

编译这本《试验统计手册》的基本立足点，是本着“古为今用”、“洋为中用”的原则，去其糟粕，取其精华，从中鉴别和借鉴有益的经验，分析和吸收有用的计算方法，以期达到对客观现象数量方面的正确认识。社会主义建设离不开对国民经济各部门的统计工作。为改进和完善统计工作的科学方法，进一步开展统计基本理论与基本方法的研究。我主要参考了 Snedecor G · W · :《Statistical Methods》; Ostile B.:《Statistical in Research》; Wilks S · S:《Mathematical Statistics》; Kendall M · G · and Buckland W · R · :《A Dictionary of Statistical Terms》以及 Owen D · B ·《Handbook of Statistical Tables》等专著编译出这本《试验统计手册》，贡献给从事专业统计工作的同志。

本书的内容，划分为三个部分：

一、试验统计公式及其推导；

二、试验统计的英文词汇及常用的符号；

三、重要的统计用表。

为了使广大读者易于理解和掌握试验统计的基本原理，在公式推导方面，主要运用代数法，仅有少数的公式采用微分法。由于试验统计问题涉及的范围广泛，因此，书中也汇集了一些与其它学科相交叉的词汇。

本书承蒙陈允蒙同志提出了宝贵的意见，深表谢意。由于编泽水平有限，书中的缺点和不足之处在所难免，敬希广大读者批评指正。

刘醒亚

目　录

第一部分　试验统计公式及其推导

一、算术平均数

$$\bar{x}=\frac{x_1+x_2+x_3+\cdots\cdots+x_n}{n} \tag{1}$$

式(1)可简写为:

$$\bar{x}=\frac{\sum_1^n x_n}{n}=\frac{\sum x}{n} \tag{2}$$

$$\bar{x}=\frac{f_1x_1+f_2x_2+\cdots\cdots f_px_p}{f_1+f_2+\cdots\cdots+f_p}=\frac{\sum fx}{n} \tag{3}$$

式中,x 为各组变量或组中点值; f 为各组的次数; p 为组数, n 为总次数。

$$\bar{x}=\overline{x'}+\frac{\sum fd'}{n}\times ci \tag{4}$$

式中, $\overline{x'}$ 为假定平均数, f 为各组的次数, d' 为各组中点与假定平均数以组距为单位的等级差; fd' 为各组次数与等级差的乘积。$\sum fd'$ 为各组次数与等级差的乘积的总和, ci 为组距。

公式推导:

$$\bar{x}=\frac{\sum fx}{n}=\frac{f_1x_1}{n}+\frac{f_2x_2}{n}+\cdots\cdots+\frac{f_nx_n}{n}$$

$$=\frac{f_1(x_1-\overline{x'}+\overline{x'})}{n}+\frac{f_2(x_2-\overline{x'}+\overline{x'})}{n}+\cdots$$

$$+\frac{f_n(x_n-\overline{x'}+\overline{x'})}{n}$$

$$=\left[\frac{f_1(x_1-\overline{x'})}{n}+\frac{f_1\overline{x'}}{n}\right]+\left[\frac{f_2(x_2-\overline{x'})}{n}+\frac{f_2\overline{x'}}{n}\right]+$$

$$\cdots+\left[\frac{f_n(x_n-\overline{x'})}{n}+\frac{f_n\overline{x'}}{n}\right]$$

$$=\left(\frac{f_1d_1}{n}+\frac{f_2d_2}{n}+\cdots+\frac{f_nd_n}{n}\right)+\frac{\sum f\overline{x'}}{n}$$

$\because \sum f=n$， $d=d'\times ci$

$$\therefore \overline{x}=\frac{\sum fd}{n}+\overline{x'}=\overline{x'}+\frac{\sum fd'}{n}\times ci$$

二、标 准 差

(一)未分组的资料的计算方法

$$S=\sqrt{\frac{\sum(x-\overline{x})^2}{n-1}} \tag{5}$$

$$S=\sqrt{\frac{\sum x^2-\frac{(\sum x)^2}{n}}{n-1}} \tag{6}$$

公式推导:

$$S=\sqrt{\frac{\sum(x-\overline{x})^2}{n-1}}=\sqrt{\frac{\sum x^2-\frac{(\sum x)^2}{n}}{n-1}}$$

$$\sum(x-\overline{x})^2=\sum x^2-2\overline{x}\sum x+n\overline{x}^2$$

$\because \sum x=n\overline{x}$

$$\therefore \sum(x-\bar{x})^2=\sum x^2-2\bar{x}n\bar{x}+n\bar{x}^2$$

$$=\sum x^2-n\bar{x}^2=\sum x^2-n\left(\frac{\sum x}{n}\right)^2$$

$$=\sum x^2-\frac{(\sum x)^2}{n}$$

式(5)可改写为:

$$S=\sqrt{\frac{\sum x^2-\frac{(\sum x)^2}{n}}{n-1}}$$

(二)归组资料的计算方法

1. 加权法

$$S=\sqrt{\frac{\sum f(x-\bar{x})^2}{n}} \tag{7}$$

$$\because \sum f(x-\bar{x})^2=\sum f(x^2-2x\bar{x}+\bar{x}^2)$$

$$=\sum fx^2-2\bar{x}\cdot\sum fx+\sum f\bar{x}^2$$

$$=\sum fx^2-2\bar{x}\cdot\sum fx+n\bar{x}^2$$

$$\because \sum fx=n\bar{x}$$

$$\therefore \sum f(x-\bar{x})^2=\sum fx^2-2\bar{x}\cdot n\bar{x}+n\bar{x}^2$$

$$=\sum fx^2-2n\bar{x}^2+n\bar{x}^2$$

$$=\sum fx^2-n\bar{x}^2$$

∴公式(7)可改写为:

$$S=\sqrt{\frac{\sum fx^2}{n}-\bar{x}^2} \tag{8}$$

2. 等级差法

$$S=\frac{\sum fd'^2}{n}-\left(\frac{\sum fd'}{n}\right)\times ci \tag{9}$$

$ci=$ 组距

$n=$ 总次数

$$d'=\frac{(x-\overline{x'})}{ci}=\frac{d}{ci}$$

公式推导:

$$S=\sqrt{\frac{\sum f(x-\overline{x})^2}{n}}$$

$$=\sqrt{\left(\frac{\sum fd'^2}{n}\right)-\left(\frac{\sum fd'}{n}\right)}\times ci$$

$$\because \sum f(x-\overline{x})^2=\sum f\,[(x-\overline{x'})-(\overline{x}-\overline{x'})]^2$$

$$=\sum f\,[(x-\overline{x'})^2-2(x-\overline{x'})(\overline{x}-\overline{x'})+(\overline{x}-\overline{x'})^2]$$

$$=\sum f(x-\overline{x'})^2-2(\sum fx-\sum f\overline{x'})(\overline{x}-\overline{x'})+\sum f(\overline{x}-\overline{x'})^2$$

$$\because \sum fx=n\overline{x}$$

$$\therefore \sum f(x-\overline{x})^2=\sum f(x-\overline{x'})^2-2n(\overline{x}-\overline{x'})^2+n(\overline{x}-\overline{x'})^2$$

$$=\sum f(x-\overline{x'})^2-n(\overline{x}-\overline{x'})^2$$

$$\because n(\overline{x}-\overline{x'})^2=n(\frac{\sum fx}{n}-\overline{x'})^2$$

$$\because \overline{x}=\frac{\sum fx}{n}$$

$$\therefore n(\frac{\sum fx}{n}-\overline{x'})^2=n\frac{(\sum fx-\sum f\overline{x'})^2}{n}$$

$$=\frac{(\sum fx-\sum f\overline{x'})^2}{n}=\frac{[\sum f(x-\overline{x'})]^2}{n}$$

$$\therefore S=\sqrt{\frac{\sum f(x-\overline{x'})^2-\frac{[\sum f(x-\overline{x'})]^2}{n}}{n}}$$

$$=\sqrt{\frac{\sum f\left(\frac{x-\overline{x'}}{ci}\right)^2-\frac{[\sum f\frac{(x-\overline{x'})}{ci}]^2}{n}}{n}}\times ci$$

$$=\sqrt{\frac{[\sum fd'^2-\frac{\sum fd'^2}{n}]}{n}}\times ci$$

$$=\sqrt{\frac{\sum fd'^2}{n}-\frac{(\sum fd')^2}{n}}\times ci$$

三、平均数标准差

$$S\overline{x}=\sqrt{\frac{S^2}{n}}=\frac{S}{\sqrt{n}} \tag{10}$$

公式推导:

设由群体抽出 K 个样本，则有 K 个$\overline{x}$，其均数为 M，则样本均数的标准差为:

$$\sigma\overline{x}=\sqrt{\frac{\sum_{1}^{K}(x-M)^2}{K}} \tag{11}$$

又每个样本有n个个体（变数），分别

用x_1，x_2，x_3……x_n表示。

用x表示任何一个。

$\overline{x}$ 与 M 差异用 d 表示，d_1d_2，d_2d_3，d_1d_3……用 d_id_j 表示。

第一个样本均数$\overline{x_1}$与差异的平方可写为:

$$(\overline{x}-M)^2=\left(\frac{x_1+x_2+x_3+\cdots+x_n}{n}-M\right)^2$$

$$=\left(\frac{x_1+x_2+x_3+\cdots+x_n-nM}{n}\right)^2$$

$$=\left[\frac{(x_1-M)+(x_2-M)+(x_3-M)+\cdots+(x_n-M)}{n}\right]^2$$

$$=\left(\frac{d_1+d_2+d_3+\cdots+d_n}{n}\right)^2$$

$$=\frac{d_1^2+d_2^2+d_3^2+\cdots+dn^2+2d_1d_2+2d_1d_3+2d_2d_3+\cdots}{n^2}$$

$$+\frac{2d_{n-1}d_n}{n^2}=\frac{\sum_1^n d^2+2\sum_1^n d_id_j}{n^2}$$

第二个样本均数$\overline{x_2}$与差异的平方同样可写为:

$$(\overline{x_2}-M)^2=\frac{\sum_1^n d^2+2\sum_1^n d_id_j}{n^2}$$

则K个样本均数$\overline{x_K}$与差异的平方为

$$(\overline{x_K}-M)^2=\frac{\sum_1^n d^2+2\sum_1^n d_id_j}{n^2}$$

将各个$(\overline{x}-M)^2$积加，则为:

$$\sum_{1}^{K}(\overline{x}-M)^2=\sum_{1}^{K}\left(\frac{\sum_{1}^{n}d^2+2\sum_{1}^{n}d_id_j}{n^2}\right)$$

$$=\sum_{1}^{K}\frac{\sum_{1}^{n}d^2}{n^2}+\sum_{1}^{K}\frac{2\sum_{1}^{n}d_id_j}{n^2}$$

d_i代表x与M差异。

d_j代表另一个x与M的差异。

而$\sum(x-\overline{x})=0$

即$\overline{x}=\dfrac{x_1+x_2+\cdots+x_n}{n}$

$\therefore(x_1+x_2+\cdots+x_n)=n\overline{x}$

$(x_1+x_2+\cdots+x_n)-n\overline{x}=0$

$(x_1-\overline{x})+(x_2-\overline{x})+\cdots+(x_n-\overline{x})=0$

$\therefore\sum(x-\overline{x})=0$

$$\therefore\sum_{1}^{K}\frac{2\sum_{1}^{n}d_id_j}{n^2}=0$$

$$\sum_{1}^{K}(\overline{x}-M)^2=\sum_{1}^{K}\sum_{1}^{n}d^2/n^2\quad=\sum_{1}^{K}\sum_{1}^{n}\frac{d^2}{n}\cdot\frac{1}{n}$$

$$\because\sigma^2=\frac{\sum(x-M)^2}{n}=\frac{\sum d^2}{n}$$

$$\therefore\sum_{1}^{K}(\overline{x}-M)^2=\sum_{1}^{K}\sigma^2\cdot\frac{1}{n}=K\frac{\sigma^2}{n}$$

$$\sigma x=\sqrt{\frac{\sum_{1}^{K}(\bar{x}-M)^2}{K}}=\sqrt{\frac{K\frac{\sigma^2}{n}}{K}}$$

$$=\sqrt{\frac{\sigma^2}{n}}=\frac{\sigma}{\sqrt{n}}$$

$$\therefore S\bar{x}=\frac{S}{\sqrt{n}}$$

式中，$S\bar{x}$为估计值。

四、t 测 验

(一)样本均数与总体均数差异显著性测验

$$t=\frac{\bar{x}-u_0}{S\bar{x}} \tag{12}$$

式中,$\bar{x}$为样本均数; u_0为总体均数。

$$df=n-1,\quad S\bar{x}=\frac{S}{\sqrt{n}}$$

(二)配对试验时的两个样本均数差异显著性测验

$$t=\frac{\bar{d}}{S\bar{d}}=\frac{\overline{x^1}-\overline{x^2}}{S\bar{d}} \tag{13}$$

$$df=n-1$$

$$S\bar{d}=\frac{\sqrt{\frac{\sum(d-\bar{d})^2}{n-1}}}{\sqrt{n}}=\sqrt{\frac{\sum(d-\bar{d})^2}{n(n-1)}}$$

式中，d 为第一、第二两个样本各对变数之差，即 x_1-x_2； $\bar{d}$ 为第一、第二两个样本各对变数平均数之差，即 x_1-x_2。

(三)非配对试验时的两个样本平均数差异显著性测验

$$t = \frac{\overline{x_1} - \overline{x_2}}{S\overline{x_1} - \overline{x_2}} \tag{14}$$

式中，$S\overline{x_1} - \overline{x_2}$ 为均数差数标准差；$\overline{x_1}$ 为第一个样本的均数；$\overline{x_2}$ 为第二个样本的均数。

$$S\overline{x_1} - \overline{x_2} = \sqrt{\frac{\sum(x_1 - \overline{x_1})^2 + \sum(x_2 - \overline{x_2})^2}{(n_1 - 1) + (n_2 - 1)}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

式中，n_1 为第一个样本（处理）的变数个数；n_2 为第二个样本（处理）的变数个数。

公式推导：

如果两个总体各作正态分布，则其样本均数的差数（$\overline{x_1} - \overline{x_2}$）必然准确地遵循正态分布即 $u\overline{x_1} - \overline{x_2} = u_1 - u_2$。

由此，两个独立样本平均数差数分布方差必等于两个总体的样本平均数方差总和。

$$\sigma^2(\overline{x_1} - \overline{x_2}) = \sigma^2\overline{x_1} + \sigma^2\overline{x_2} = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}$$

$$\sigma(\overline{x_1} - \overline{x_2}) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

将σ改写为S，则$S\overline{x_1} - \overline{x_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$

当$n_1 = n_2 = n$时：

则
$$\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$
$$= \sqrt{\frac{S_1^2 + S_2^2}{n}}$$

$$=\sqrt{\frac{1}{2}\left[2\left(\frac{S_1^2+S_2^2}{n}\right)\right]}$$

$$=\sqrt{\frac{1}{2}\left(\frac{S_1^2+S_2^2}{n_1}+S_1^2+S_2^2\right)n_2}$$

$$=\sqrt{\frac{1}{2}(S_1^2+S_2^2)\left(\frac{1}{n_2}+\frac{1}{n_2}\right)}$$

又$S_1=\sqrt{\frac{\sum(x_1-\overline{x_1})^2}{n-1}}$ $\quad S_1^2=\frac{\sum(x_1-\overline{x_1})^2}{n-1}$

$S_2=\sqrt{\frac{\sum(x_2-\overline{x_2})^2}{n-1}}$ $\quad S_2^2=\frac{\sum(x_2-\overline{x_2})^2}{n-1}$

代入上式得:

$S\overline{x_1}-\overline{x_2}$

$$=\sqrt{\frac{\sum(x_1-\overline{x_1})^2+\sum(x_2-\overline{x_2})^2}{n_1+n^2-2}\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

又因$\sum(x_1-\overline{x_1})^2=\sum_1^{X2}-\frac{(\sum x_1)^2}{n_1}$

同理$\sum(x_2-\overline{x_2})^2=\sum_2^{X2}-\frac{(\sum x_2)^2}{n_2}$

(四)样本百分数与总体百分数的差异比较显著性测验

$$t=\frac{dp}{\sigma p} \tag{15}$$

式中,$\sigma p=\sqrt{\frac{pq}{n}}$; n为总体个体数;p为总体百分数。

$p+q=1 \qquad q=1-p$

$\sigma^2=\frac{\sum f(x-u)^2}{n}=np(1-p)/n \quad =p(1-p)$

$\therefore\sigma^2=pq \qquad \sigma=\sqrt{pq}$

又$\sigma^2\bar{x}=\dfrac{\sigma^2}{n}=\dfrac{pq}{n}$

$\therefore\sigma\bar{x}(\sigma p)=\sqrt{\dfrac{pq}{n}}$

(五)两个百分数样本差异显著性测验

$$t=\frac{\hat{p}_1-\hat{p}_2}{\sigma\hat{p}_1-\hat{p}_2} \tag{16}$$

式中，$\hat{p}_1$ 为第一个样本发生的百分数$\dfrac{x_1}{n_1}\times100$；$\hat{p}_2$ 为第二个样本发生的百分数$\dfrac{x_2}{n_2}\times100$；$\sigma\hat{p}_1-\hat{p}_2$ 为两样本百分数间差异的标准差；$\bar{p}$ 表示两个样本共同发生的百分数（加权平均值）即$\dfrac{x_1+x_2}{n_1+n_2}$。

又$\bar{q}=1-\bar{p}$　　$\sigma p_1=\sqrt{\dfrac{\bar{p}\bar{q}}{n_1}}$

$\sigma p_2=\sqrt{\dfrac{\bar{p}\bar{q}}{n_2}}$

$$\begin{aligned}\therefore\sigma\hat{p}_1-\hat{p}_2&=\sqrt{\frac{\bar{p}\bar{q}}{n_1}+\frac{\bar{p}\bar{q}}{n_2}}\\&=\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\\&=\sqrt{\frac{x_1+x_2}{n_1+n_2}\left(1-\frac{x_1+x_2}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\end{aligned}$$

五、X²(卡方)检验

(一)适合性检验

比较实际数与理论假设是否相符合的检验，在试验统