数理统计题解

全国高等农林专科统编教材

毕庆雨 主编

中国林业出版社

全国高等农林专科统编教材

数理统计题解

毕庆雨 主编

林业专业用

中国林业出版社

全国高等农林专科统编教材

数理统计题解

毕庆雨　主编

中国林业出版社出版（北京西城区刘海胡同7号）
新华书店北京发行所发行　昌黎县印刷厂印刷

787×1092毫米　16开本　18.5印张　365千字
1994年5月第一版　1994年5月第一次印刷
印数1—4300册　　定价：7.20元
ISBN 7-5038-1128-5/S·0632

(京)新登字033号

主　编　毕庆雨(河北林学院)
副主编　陈华豪(东北林业大学)
编　者　林新荣(内蒙古林学院)
审稿人　李世达(主审　东北林业大学)
　　　　黄用廉(北京林业大学)
责任编辑　刘先银

前 言

本书是为全国高等农林专科林业专业数理统计课程编写的参考教材。它是在国家教委高教司和全国高等农林专科基础课程教材委员会的直接领导和组织下编写的。

继《数理统计》一书出版以后，为了使教材进一步注意体现“突出应用性，加强实践性，强调针对性”的特点，使读者更好地学习数理统计这门课，我们又编写了《数理统计题解》这本书。

为了使读者在学习数理统计这门课程中，学到更多的结合林业生产和林业科学研究的实践知识，能够比较熟练地应用数理统计的基本方法解决林业当中的实际问题。在编写过程中我们参阅了各农林院校编写的数理统计、生物统计及有关林业专业的教材。在例题与习题的选取上，着重结合林业的实际问题，并照顾到林业各专业的需要。

为了扩大学生的知识面，提高学生解决实际问题的能力，题解中部分例题与习题已超出专科数理统计教学大纲的范围，所以本题解也可作为林业本科生参考用书。在编写过程中力求做到通俗易懂，由浅入深，理论联系实际，便于自学。所以本书也是成人教育的重要参考教材。

本书主要内容有参数估计，假设检验，方差分析，回归分析和试验设计。每章概括了基本内容并给出了主要公式，列举了各种类型的例题，选择了部分习题，并附有习题答案供参考。学习本书要求读者须具备概率论的基本知识，掌握数理统计的基本方法。

本书在编写过程中，得到国家教委高教司，林业部，全国高等农林专科基础课程教材委员会，河北林学院，东北林业大学，内蒙古林学院，北京林业大学的领导及有关同志的大力支持与帮助，在此我们表示诚挚的谢意。

由于我们的业务水平有限，专业知识不足，时间仓促，缺点错误在所难免，希望广大读者提出宝贵意见，以便今后改正。

编 者

1992年8月

目　　录

第一章　参数估计…………（1）

基本内容…………（1）

一、总体与样本…………（1）

二、参数估计…………（2）

例题…………（6）

习题一…………（32）

第二章　统计假设检验…………（35）

基本内容…………（35）

一、正态总体平均数的假设检验…………（35）

二、两个样本平均数间的差异显著性检验…………（36）

三、总体频率的假设检验…………（36）

四、样本频率间的差异显著性检验…………（37）

五、方差的齐性检验…………（37）

六、分布的假设检验…………（38）

七、适合性检验与独立性检验…………（38）

例题…………（38）

习题二…………（68）

第三章　方差分析…………（73）

基本内容…………（73）

一、单因素的方差分析…………（73）

二、双因素的方差分析…………（74）

例题…………（76）

习题三…………（127）

第四章　回归分析…………（136）

基本内容…………（136）

一、一元线性回归…………（136）

二、非线性回归…………（137）

三、多元线性回归…………（138）

四、逐步回归…………（140）

五、数量化回归…………（142）

例题…………（144）

习题四…………（189）

第五章 试验设计…………………………………………………………………… (193)
基本内容……………………………………………………………………………… (193)
一、试验设计的基本原则……………………………………………………………… (193)
二、随机区组试验……………………………………………………………………… (193)
三、拉丁方试验………………………………………………………………………… (193)
四、平衡不完全区组试验……………………………………………………………… (194)
五、正交试验法………………………………………………………………………… (194)
例题…………………………………………………………………………………… (194)
习题五………………………………………………………………………………… (233)
习题答案……………………………………………………………………………… (237)
附表…………………………………………………………………………………… (241)
1. 正态分布的密度函数表 ………………………………………………………… (241)
2. 正态分布表 ……………………………………………………………………… (242)
3. 正态分布的双侧分位数 (u_a) 表………………………………………………… (244)
4. 二项分布参数 p 的置信区间表 ………………………………………………… (245)
5. 泊松(Poisson)分布表…………………………………………………………… (249)
6. 泊松(Poisson)分布参数 λ 的置信区间表………………………………… (255)
7. χ^2 分布的上侧分位数(χ^2_a)表………………………………………… (256)
8. 学生氏 t 分布表 ……………………………………………………………… (257)
9. 学生氏 t 分布的双侧分位数 (t_a) 表 ……………………………………… (259)
10. F 检验的临界值 (F_a) 表……………………………………………………… (260)
11. 随机数表………………………………………………………………………… (265)
12. 多重比较中的 q 表……………………………………………………………… (267)
13. 多重比较中的 S 表……………………………………………………………… (269)
14. 检验相关系数 $\rho=0$ 的临界值 (r_a) 表……………………………………… (270)
15. r 与 z 的换算表 ……………………………………………………………… (270)
16. Duncan's 新复极差测验 a 为 0.05 及 0.01 时的 SSR 值表 ………………… (271)
17. 平衡不完全区组设计表………………………………………………………… (273)
18. 常用正交表……………………………………………………………………… (282)
参考文献……………………………………………………………………………… (289)
全国高等农林专科基础课程第一批统编教材书目…………………………………… (290)

第一章　参数估计

基本内容

一、总体与样本

1、总体　按着统计研究目的而确定的同类事物或现象的全体称为总体。组成总体的每个单元称为总体单元。总体中含有的单元数，称为总体容量。

表 1-1

特征数	总体	样本
平均数	$\mu=\overline{X}=E(\xi)=\frac{1}{N}\sum_{i=1}^{N}X_i$	$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ $\bar{x}=\frac{1}{n}\sum_{i=1}^{k}x_i f_i$，$f_i$为$x_i$的频数
总量	$T=\sum_{i=1}^{N}X_i=N\overline{X}$	$T_s=\sum_{i=1}^{n}x_i=n\bar{x}$
方差	$\sigma^2=\frac{1}{N}\sum_{i=1}^{N}(X_i-\overline{X})^2$	$s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ $s^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2$ $s^2=\frac{1}{n}\sum_{i=1}^{k}(x_i-\bar{x})^2 f_i$ $=\frac{1}{n}\sum_{i=1}^{k}x_i^2 f_i-\left(\sum_{i=1}^{k}x_i f_i/n\right)^2$
标准差	$\sigma=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i-\overline{X})^2}$	$s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ $s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2}$ $s=\sqrt{\frac{1}{n}\sum_{i=1}^{k}(x_i-\bar{x})^2 f_i}$ $=\sqrt{\frac{1}{n}\sum_{i=1}^{k}x_i^2 f_i-\left(\sum_{i=1}^{k}x_i f_i/n\right)^2}$
变异系数	$V_x=\frac{\sigma}{\mu}$	$v_x=\frac{s}{\bar{x}}$
频率	$W=\frac{M}{N}$	$w=\frac{m}{n}$
极差	$D=X_{\max}-X_{\min}$	$d=x_{\max}-x_{\min}$

例如，当我们调查某林区林木的平均树高时，该林区的全体林木便构成一个总体，其中每一株树的树高就是一个单元。但应注意，总体单元未必是总体的自然单元，也可以规定以一定面积(如 0.04 ha)上的蓄积量作为一个总体单元。

从前面的例子可以看出，在数理统计的实际问题中，我们关心的是研究对象的某个数量指标。前面提到的每株树的树高，一定面积上的蓄积量，尽管在相同的条件下，由于各种随机因素的影响，每一株树的树高，每块一定面积(如 0.04 ha)上的蓄积量不尽相同。因此，这类数量指标是随机变量 ξ，ξ 取遍数量指标的全体数值，恰好就是总体。所以常用随机变量 ξ 来代表总体，而 ξ 的每一个可能值就是总体单元。

2．样本　由全部总体单元中随机抽取一部分单元，所抽取的这一部分单元称为样本。样本中的单元称为样本单元，样本中含有的单元数称为样本容量。

3．重复抽样与不重复抽样　采取有放回的方式抽取样本称为重复抽样，采取无放回的方式抽取样本称为不重复抽样。

4．总体特征数与样本特征数(表 1-1)

二、参 数 估 计

1．总体平均数的抽样估计

(1) 大样本重复抽样的估计方法　样本平均数 $\bar{x}$ 作为总体平均数 μ 的估计值，误差限为

$$\Delta(\bar{x})=\frac{u_\alpha \sigma}{\sqrt{n}}\quad(\text{当 } \sigma^2 \text{ 已知})$$

或

$$\Delta(\bar{x})\approx\frac{u_\alpha s}{\sqrt{n-1}}\quad(\text{当 } \sigma^2 \text{ 未知})$$

相对误差限为

$$\Delta'(\bar{x})=\frac{\Delta(\bar{x})}{\mu}\approx\frac{\Delta(\bar{x})}{\bar{x}},\quad \Delta(\bar{x})<\mu。$$

估计精度为

$$P_c(\bar{x})=1-\Delta'(\bar{x}),\quad \Delta(\bar{x})<\mu。$$

可靠性为 $P=1-\alpha$。

总体平均数 μ 的 $P=1-\alpha$ 的置信区间为

$$\left[\left(\bar{x}-\frac{u_\alpha \sigma}{\sqrt{n}}\right),\ \left(\bar{x}+\frac{u_\alpha \sigma}{\sqrt{n}}\right)\right]\quad(\text{当 } \sigma^2 \text{ 已知})$$

或

$$\left[\left(\bar{x}-\frac{u_\alpha s}{\sqrt{n-1}}\right),\ \left(\bar{x}+\frac{u_\alpha s}{\sqrt{n-1}}\right)\right]\quad(\text{当 } \sigma^2 \text{ 未知})$$

必要抽样数即样本单元数为

$$n_0=\left[\frac{u_\alpha\sigma}{\Delta(\bar{x})}\right]^2\approx\left[\frac{u_\alpha s}{\Delta(\bar{x})}\right]^2$$

或

$$n_0=\left[\frac{u_\alpha\sigma}{\mu\Delta'(\bar{x})}\right]^2\approx\left[\frac{u_\alpha s}{\bar{x}\Delta'(\bar{x})}\right]^2$$

$$n_0=\left[\frac{u_\alpha V_x}{\Delta'(\bar{x})}\right]^2\approx\left[\frac{u_\alpha v_x}{\Delta'(\bar{x})}\right]^2$$

（2）大样本不重复抽样的估计方法　样本平均数 $\bar{x}$ 作为总体平均数 μ 的估计值，误差限为

$$\Delta(\bar{x})=\frac{u_\alpha\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\ （当 \sigma^2 已知）$$

或

$$\Delta(\bar{x})\approx\frac{u_\alpha s}{\sqrt{n-1}}\sqrt{1-\frac{n}{N}}\ （当 \sigma^2 未知）$$

可靠性为$P=1-\alpha$。

总体平均数 μ 的 $P=1-\alpha$ 的置信区间为

$$\left[\left(\bar{x}-\frac{u_\alpha\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right),\ \left(\bar{x}+\frac{u_\alpha\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right)\right]$$

或

$$\left[\left(\bar{x}-\frac{u_\alpha s}{\sqrt{n-1}}\sqrt{1-\frac{n}{N}}\right),\ \left(\bar{x}+\frac{u_\alpha s}{\sqrt{n-1}}\sqrt{1-\frac{n}{N}}\right)\right]$$

样本单元数为

$$n_1=\frac{u_\alpha^2\sigma^2N}{\Delta^2(\bar{x})(N-1)+u_\alpha^2\sigma^2}$$

或

$$n_1=\frac{n_0}{1+\frac{n_0}{N}}$$

（3）正态总体平均数重复抽样的估计方法用样本平均数 $\bar{x}$ 作为总体平均数 μ 的估计值，误差限为

$$\Delta(\bar{x})=\frac{t_\alpha s}{\sqrt{n-1}}$$

相对误差限为

$$\Delta'(\bar{x})=\frac{\Delta(\bar{x})}{\mu}\approx\frac{\Delta(\bar{x})}{\bar{x}}$$

估计精度为

$$P_c(\bar{x})=1-\Delta'(\bar{x})\approx 1-\frac{\Delta(\bar{x})}{\bar{x}}$$

可靠性为$P=1-\alpha$

总体平均数μ的$P=1-\alpha$的置信区间为

$$\left[\left(\bar{x}-\frac{t_\alpha s}{\sqrt{n-1}}\right),\ \left(\bar{x}+\frac{t_\alpha s}{\sqrt{n-1}}\right)\right]$$

2．总体频率的抽样估计（大样本）

(1) 重复抽样的估计方法

① 用正态分布估计总体频率的方法　用样本频率w作为总体频率W的估计值，误差限为

$$\Delta(w)=u_\alpha\sqrt{\frac{W(1-W)}{n}}$$

相对误差限为

$$\Delta'(w)=\frac{\Delta(w)}{W}\approx\frac{\Delta(w)}{w}$$

估计精度为

$$P_c(w)=1-\Delta'(w)$$

可靠性为$P=1-\alpha$。

总体频率W的$1-\alpha$的置信区间为

$$\left[\left(w-u_\alpha\sqrt{\frac{W(1-W)}{n}}\right),\ \left(w+u_\alpha\sqrt{\frac{W(1-W)}{n}}\right)\right]$$

但我们看到不论是误差限$\Delta(w)$还是置信区间的上、下限，在计算时都要用未知参数W之值。但是当n充分大时，可用w近似代替W。也可用概率式子

$$P\left\{\frac{n}{n+u_\alpha^2}\left(w+\frac{u_\alpha^2}{2n}-u_\alpha\sqrt{\frac{w(1-w)}{n}+\frac{u_\alpha^2}{4n^2}}\right)\leqslant W\leqslant\right.$$

$$\left.\frac{n}{n+u_\alpha^2}\left(w+\frac{u_\alpha^2}{2n}+u_\alpha\sqrt{\frac{w(1-w)}{n}+\frac{u_\alpha^2}{4n^2}}\right)\right\}=1-\alpha$$

求得W的$P=1-\alpha$的精确的置信区间。

样本单元数

$$n_0=\frac{u_\alpha^2W(1-W)}{\Delta^2(w)}\approx\frac{u_\alpha^2w(1-w)}{\Delta^2(w)}$$

或

$$n_0=\frac{u_\alpha^2(1-W)}{W\Delta'^2(w)}\approx\frac{u_\alpha^2(1-w)}{w\Delta'^2(w)}$$

② 用泊松分布估计总体频率的方法　设c为随机变量m在抽样结果中取得的数值，则当n充分大，W很小时，有

$$P=\left\{w=\frac{c}{n}\right\}=\frac{\lambda^{c}e^{-\lambda}}{c\,!}$$

即样本频率近似地服从参数为 $\lambda=nW$ 的泊松分布。可由泊松分布参数 λ 的置信区间表中查到 λ 的置信区间，再由 $W=\frac{\lambda}{n}$ 可得到W的置信区间。

(2). 不重复抽样的估计方法 以w在抽样结果中所取得的数值作为W的估计值，误差限为

$$\Delta(w)=u_{\alpha}\sqrt{\frac{W(1-W)}{n}\left(\frac{N-n}{N-1}\right)}$$

可靠性为$P=1-\alpha$。

总体频率W的$P=1-\alpha$的置信区间为

$$\left[\left(w-u_{\alpha}\sqrt{\frac{W(1-W)}{n}\left(\frac{N-n}{N-1}\right)}\right),\ \left(w+u_{\alpha}\sqrt{\frac{W(1-W)}{n}\left(\frac{N-n}{N-1}\right)}\right)\right]$$

当n充分大时，可用w近似地代替 $\Delta(w)$ 中及置信区间上下限中的W，也可用概率式子

$$P\left\{\frac{n}{n+u_{\alpha}^{2}\left(\frac{N-n}{N-1}\right)}\left[w+\frac{u_{\alpha}^{2}}{2n}\left(\frac{N-n}{N-1}\right)\right.\right.$$

$$\left.-u_{\alpha}\sqrt{\frac{w(1-w)(N-n)}{n(N-1)}+\frac{u_{\alpha}^{2}}{4n^{2}}\left(\frac{N-n}{N-1}\right)^{2}}\ \right]\leqslant$$

$$W\leqslant\frac{n}{n+u_{\alpha}^{2}\left(\frac{N-n}{N-1}\right)}\left[w+\frac{u_{\alpha}^{2}}{2n}\left(\frac{N-n}{N-1}\right)+\right.$$

$$\left.\left.u_{\alpha}\sqrt{\frac{w(1-w)(N-n)}{n(N-1)}+\frac{u_{\alpha}^{2}}{4n^{2}}\left(\frac{N-n}{N-1}\right)^{2}}\ \right]\right\}=1-\alpha$$

求得W的$P=1-\alpha$的精确的置信区间。

样本单元数

$$n_{1}=\frac{u_{\alpha}^{2}W(1-W)N}{\Delta^{2}(w)(N-1)+u_{\alpha}^{2}W(1-W)}$$

或者，在N充分大时，$\frac{N-1}{N}$可近似地用1代替，即得

$$n_{1}=\frac{n_{0}}{1+\frac{n_{0}}{N}}$$

3. 总体频率的抽样估计（小样本）

设总体N个单元中有M个单元具有某种特点，则总体频率$W=\frac{M}{N}$。利用等概重复抽样方法，从总体中抽取n个单元组成样本，有m个单元具有某种特点，m是服从二项分布的

随机变量，参数为n，W。以k表示m在抽样结果中所取得的数值，则由于$w=\frac{m}{n}$，所以

$$P\left\{w=\frac{k}{n}\right\}=P\{m=k\}=C_n^k W^k(1-W)^{n-k}$$

按照样本中具有某种特点单元的个数k和不具有某种特点单元的个数$n-k$，查附表（二项分布参数p的置信区间表），所得两个数字就构成总体频率W的置信区间。

4．总体平均数的分层抽样估计方法

（1）总体平均数的估计值为

$$\bar{y}_{st}=\sum_{h=1}^{L} W_h \bar{y}_h$$

（2）估计的误差限为

① 大样本情形

$$\Delta(\bar{y}_{st})=u_\alpha\sqrt{\sum_{h=1}^{L}\frac{W_h^2 s_h^2}{n_h-1}} \qquad \text{（重复抽样）}$$

$$\Delta(\bar{y}_{st})=u_\alpha\sqrt{\sum_{h=1}^{L}\frac{W_h^2 s_h^2}{n_h-1}\left(1-\frac{n_h}{N_h}\right)} \qquad \text{（不重复抽样）}$$

② 小样本情形

$$\Delta(\bar{y}_{st})=t_\alpha\sqrt{\frac{1}{n(n-L)}\sum_{h=1}^{L} n_h s_h^2}$$

t_α是与概率α相对应的t分布双侧分位数。自由度$f=n-L$。

（3）估计精度为

$$P_c(\bar{y}_{st})=1-\frac{\Delta(\bar{y}_{st})}{\bar{y}_{st}}$$

例　　题

例 1.1　由某种苗木组成的总体中，用重复抽样方式随机抽取 50 株苗木组成样本，测得样本中各苗木的地径资料如下(单位：cm)：

2.65，3.24，2.80，3.00，2.74，2.12，3.58，3.60，3.52，2.24，2.67，2.57，
3.00，1.64，2.97，2.98，3.51，3.54，2.73，3.53，2.04，2.53，2.98，3.15，
3.30，2.58，2.59，1.65，1.76，2.23，3.40，2.05，2.63，2.48，2.31，3.42，
3.21，2.70，1.14，2.59，2.55，1.35，1.20，3.75，3.80，4.00，2.45，2.98，
2.10，2.04。

（1）试求该样本在地径标志上的样本平均数，样本标准差，样本极差，样本变异系数及地径不小于 2 cm 的样本频率。

（2）试将样本分组整理，列出样本频率分布表。

解　（1）样本平均数

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i=\frac{1}{50}(2.65+3.24+\cdots+2.04)$$

$$=\frac{135.59}{50}=2.7118$$

样本标准差

$$s=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2}=\sqrt{\frac{1}{50}\times 391.0357-2.7118^2}$$

$$=0.6832$$

样本极差

$$d=x_{\max}-x_{\min}=4.00-1.14=2.86$$

样本变异系数

$$v_x=\frac{s}{\bar{x}}=\frac{0.6832}{2.7118}=0.2519=25.19\%$$

地径不小于 2 cm 的样本频率

$$w=\frac{m}{n}=\frac{44}{50}=0.88$$

（2）将样本资料分成 9 组，并将样本资料所在区间[1.14，4.00]略扩成[1.03，4.09]，确定公共组距为 0.34 cm，并采取下限排外法（各组包含上限而不包含下限），列出样本频数、频率分布表(表 1-2)。

表 1-2

组区间	组中值 x_i	频数 f_i	频率 f_i/n
1.03—1.37	1.20	3	0.06
1.37—1.71	1.54	2	0.04
1.71—2.05	1.88	4	0.08
2.05—2.39	2.22	5	0.10
2.39—2.73	2.56	13	0.26
2.73—3.07	2.90	8	0.16
3.07—3.41	3.24	5	0.10
3.41—3.75	3.58	8	0.16
3.75—4.09	3.92	2	0.04

例 1.2　为了解一批种子的发芽情况，随机抽取 n 粒种子作发芽试验，测得有 m 粒发芽。试求发芽率的极大似然估计。

解　设随机变量 ξ 表示一粒种子的发芽情况，“$\xi=1$”表示种子发芽，“$\xi=0$”表示种子不发芽。于是可知，$P\{\xi=1\}=p$，$P\{\xi=0\}=q=1-p$。显然，ξ 服从(0—1)分布

$$P\{\xi=k\}=p^k q^{1-k} \qquad (k=0,\ 1)$$

所试验的 n 粒种子可以看作是服从(0—1)分布的总体的一个容量为 n 的随机样本 x_1，$x_2,\cdots$，x_n，其中有 m 粒发芽意味着满足条件 $\sum_{i=1}^{n}x_i=m$。

似然函数

$$L(p)=L(x_1,\ x_2,\ \cdots,\ x_n;\ p)$$
$$=\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}$$
$$=p^{\sum_{i=1}^{n} x_i}(1-p)^{\sum_{i=1}^{n}(1-x_i)}$$

将上式两边取对数

$$\ln L(p)=\left(\sum_{i=1}^{n} x_i\right)\ln p+\left[\sum_{i=1}^{n}(1-x_i)\right]\ln(1-p)$$

似然方程为

$$\frac{d}{dp}\ln L(p)=\frac{1}{p}\sum_{i=1}^{n} x_i-\frac{1}{1-p}\left(\sum_{i=1}^{n}(1-x_i)\right)$$
$$=\frac{1}{p}\sum_{i=1}^{n} x_i-\frac{1}{1-p}\left(n-\sum_{i=1}^{n} x_i\right)=0$$

由于$\sum_{i=1}^{n} x_i=m$，所以有

$$\frac{m}{p}-\frac{n-m}{1-p}=0$$

解此方程有

$$\hat{p}=\frac{m}{n}$$

故发芽率的极大似然估计值 $\hat{p}=\frac{m}{n}$

如果有一批种子，为了估计其发芽率，随机抽取 80 粒，进行发芽试验，结果有 76 粒发芽。则这批种子发芽率的极大似然估计值为

$$\hat{p}=\frac{m}{n}=\frac{76}{80}=0.95$$

例 1.3　随机抽取 60 个木材试件进行测定，其含水率(%) 的资料整理如表 1-3。试以 95.5%的可靠性估计该木材的平均含水率。

表 1-3

分　组	8—9	9—10	10—11	11—12	12—13	13—14	14—15	15—16	16—17
组中值 x_i	8.5	9.5	10.5	11.5	12.5	13.5	14.5	15.5	16.5
频数 f_i	4	5	8	10	12	9	7	3	2

解
$$\bar{x}=\frac{1}{n}\sum_{i=1}^{k} x_i f_i=\frac{733}{60}=12.22$$

$$s=\sqrt{\frac{1}{n}\sum_{i=1}^{k}x_i^2f_i-\bar{x}^2}=\sqrt{3.955}=1.989$$

由于 $P=0.955$，则 $\alpha=0.045$，查正态分布的双侧分位数表，有 $u_\alpha\approx2$。因此可作如下结论：

(1) 定值估计法：该木材平均含水率 μ 的估计值 $\bar{x}=12.22$，误差限 $\Delta(\bar{x})\approx\frac{u_\alpha s}{\sqrt{n-1}}$ $=\frac{2\times1.989}{\sqrt{59}}=0.518$。

(2) 区间估计法：μ 的 95.5%的置信区间为[(12.22－0.518)，(12.22＋0.518)] 即[11.702，12.738]。该木材的平均含水率在 11.702%与 12.738%之间的概率为 95.5%。

例 1.4 用等概重复抽样方式，由某块人工幼龄林地上的全部林木所组成的总体中，随机抽取 100 株林木组成样本。样本中各林木的胸径资料如下表(表 1-4)，试以 95%的可靠性，对该幼龄林地上全部林木的平均胸径进行估计。

表 1-4

胸 径 x_i(cm)	频 数 f_i	$f_i x_i$	$f_i x_i^2$
4.0	1	4.0	16.00
4.2	10	42.0	176.40
4.4	20	88.0	387.20
4.6	44	202.4	931.04
4.8	16	76.8	368.64
5.0	8	40.0	200.00
5.2	1	5.2	27.04
Σ	100	458.4	2 106.32

解 由表中计算可得

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{k}f_ix_i=\frac{1}{100}\times458.4=4.584$$

$$s=\sqrt{\frac{1}{n}\sum_{i=1}^{k}f_ix_i^2-\left(\frac{1}{n}\sum_{i=1}^{k}f_ix_i\right)^2}$$

$$=\sqrt{\frac{2106.32}{100}-\left(\frac{458.4}{100}\right)^2}=\sqrt{21.0632-21.0131}$$

$$=0.2238$$

由于 $P=0.95$，则 $\alpha=0.05$，相应 $u_\alpha=1.96$，因此可作如下结论：

(1) 定值估计法：总体平均数 μ 的估计值为 $\bar{x}=4.584$ cm，误差限为 $\Delta(\bar{x})\approx\frac{u_\alpha s}{\sqrt{n-1}}$

$=\dfrac{1.96\times 0.2238}{\sqrt{100-1}}=0.044\ \mathrm{cm}$，可靠性为95%。

或将误差限换为精度，则有

$$P_c(\bar{x})=1-\Delta'(\bar{x})\approx 1-\frac{\Delta(\bar{x})}{\bar{x}}=1-\frac{0.044}{4.584}=0.9904$$
$$=99.04\%$$

因此，结论为：$\bar{x}=4.584$ 是 μ 的估计值，精度为99.04%、可靠性为95%。

(2) 区间估计法：总体平均数 μ 的95%的置信区间为

[(4.584−0.044)，(4.584+0.044)]即[4.540，4.628]。

例 1.5 设采取等概重复抽样方式，由某块人工马尾松林地上全部林木所组成的总体中，抽取了50株林木组成样本，样本中各林木的胸径资料如下(单位：cm)：

11.7，13.6，12.0，10.2，7.0， 12.6，11.2，10.2，10.0，14.2，14.5，10.8，
7.2， 5.1， 11.6，8.5， 7.2， 4.2， 10.3，8.7， 10.5，11.7，7.6， 11.3，
6.6， 9.0， 9.4， 13.0，9.5， 10.0，9.7， 11.5，8.4， 9.5， 10.5，9.3，
8.0， 11.1，7.1， 5.9， 8.7， 12.5，8.5， 10.2，10.9，9.1， 12.2，10.6，
8.2， 10.3。

试以95%的可靠性，对该人工马尾松林的平均胸径进行估计（假定胸径满足正态分布）。

解 样本平均数 $\bar{x}=\dfrac{1}{50}\sum\limits_{i=1}^{50}x_i=9.832$，样本标准差 $s=\sqrt{\dfrac{1}{50}\sum\limits_{i=1}^{50}x_i^2-\bar{x}^2}=2.2074$。

$P=0.95$，则 $\alpha=0.05$，所以 $u_\alpha=1.96$，因此，可作结论如下：

(1) 定值估计法：总体平均数 μ 的估计值 $\bar{x}=9.832\ \mathrm{cm}$，误差限为 $\Delta(\bar{x})\approx u_\alpha\dfrac{s}{\sqrt{n-1}}$

$=1.96\times\dfrac{2.2074}{\sqrt{50-1}}=0.6181\ \mathrm{cm}$，可靠性为95%。

或将误差限换为精度，则由于

$$P_c(\bar{x})=1-\Delta'(\bar{x})\approx 1-\frac{\Delta(\bar{x})}{\bar{x}}=1-\frac{0.6181}{9.832}$$
$$=0.9371=93.71\%$$

因此，结论为：$\bar{x}=9.832$ 是 μ 的估计值，精度为93.71%，可靠性为95%。

(2) 区间估计法：μ 的95%的置信区间为[(9.832−0.6181)，(9.832+0.6181)]或[9.2139，10.4501]。即平均胸径在9.2139 cm与10.4501 cm之间的概率为95%。

例 1.6 在面积为250 ha范围的林区内，用等概重复的抽样方式，随机抽取225块面积为0.25 ha的样地组成样本，根据样地上全面测量的蓄积量资料求出每0.25 ha的平均蓄积量为22 m^3，标准差为2.5 m^3。试以95.5%的可靠性估计全林每公顷平均蓄积量。