

水文統計的 原理与方法

金光炎編

水利电力出版社

水文統計的 原理与方法

金光炎編

水利电力出版社

水文統計的原理与方法

金 光 炎編

*

1177 S 303

水利电力出版社出版 (北京西郊科学路二里内)

北京市书刊出版业营业许可证出字第105号

水利电力出版社印刷厂排印

新华书店科技发行所发行 各地新华书店經售

*

850×1168 $\frac{1}{16}$ 开本 * 11 $\frac{1}{4}$ 印張 * 303千字 * 定价(第10类)2.00元

1959年5月北京第1版

1960年3月北京第2次印刷(2,591—4,760册)

序 言

近年来，水文計算中运用数理統計法来分析水文現象日益增多，許多工程人員也迫切要求深入了解这方面的知識。但我国目前还缺少比較完整的水文統計書籍，为此，編者收集了現在常用的各項水文統計方法，編成本書。書中除闡述各家的理論和計算的方法外，並提出了一些个人的看法。

利用数理統計法的优点是可以給出水文現象的偶然性規律。大家知道，当前的气象科学，还不能对降雨作出长期定量的預报，因此也就不能确切推算某水利工程在运用期間內会必然遭遇的各种水文特征，但是利用数理統計法我們却有可能估計这些工程在运用期中所可能遭遇的各种水文情况，解决了水利工程規劃中所提出的問題。

不过，由于数理統計法的广泛应用，也使若干工程人員对这种方法产生了一些誤解。他們过高地評價了这个方法，以至形式地和机械地利用了这个方法而不去認真和客觀地分析原始資料和現象的本質，忽視了自然地理的影响，显然，这是不对的。应当指出，数理統計法的基本職能只是各种技术性的計算，其本身也帶有若干缺点，故在統計分析的同时，必須密切結合現象的物理成因分析和考虑地理分佈的綜合，这样才能产生最好的效果，否則就容易脱离实际，甚至歪曲事实。

本書共分六章，篇首簡述了統計研究的內容，其次分別詳細地敘述了水文計算中的統計参数、机率理論、頻率計算、相关分析和抽样誤差等，并对每个細節和公式都作了詳細的推演与考証。

这里还需說明一下，即在頻率計算中，对于一般的方法，如特大值处理、抽样誤差的估計和适綫示例等，仅列于皮尔逊第三型曲綫的一节中。其他綫型的处理方法大都是相同的，只要掌握

了这些方法的来源，就不难加以类推，故書中不再贅述。又数理統計的計算方法甚为簡單，因而本書对同类的問題，也不再重复举例，以节省篇幅；有些方法仅用文字上說明已能代替例子的，也都省去了。

1956年6月間，編者写成了本書的前身“水文学中的数理統計法”，并由治淮委员会勘测設計院加以油印。去年来北京水利科学研究院工作，又接受了編写本書的任务。本書初稿完成于去年9月份，現在又作了少許的修改与补充。

初稿完成前后，中国科学院数学研究所张里千先生給我們系統地介紹了机率論与数理統計方面的知識，使我們澄清了許多机率論与数理統計上的問題，对編写本書起了較大的作用，書中某些地方并引用了张先生的講課筆記(1—20)。

編者自1955年初开始接触水文統計工作以来，蒙淮委及水利科学研究院領導和同志們的鼓励，以及华东水利学院和数学研究所許多同志的帮助与启发，这是完成本書的重要因素，在此表示衷心的感謝。

本書主要是根据〔1—1〕〔2—1〕〔4—4〕〔4—5〕和水利科学研究院所收集的各有关頻率曲綫的文献編成的，可以作为水文工作者深入研究之用，也可以作为大专学校的水文专业学生及其他数理統計工作者的参考。

限于我的能力与水平，其中的缺点和錯誤一定很多，恳切地盼望着同志們的批評与指正。

金光炎写于水利科学研究院

1958.2.北京

目 錄

第一章 總 論

§1-1 統計的概念	5
§1-2 統計觀察	6
§1-3 統計歸納	9
§1-4 水文計算與数理統計	12

第二章 統計參數

§2-1 總 述	13
§2-2 平均數	14
§2-3 算術平均數	16
§2-4 中位數	21
§2-5 眾 數	23
§2-6 幾何平均數	25
§2-7 調和平均數	26
§2-8 矩	27
§2-9 离散度	31
§2-10 偏 度	36
§2-11 峯 度	38
§2-12 簡化計算與近似公式	40

第三章 機率論的基本知識

§3-1 總 述	46
§3-2 機率的觀念	47
§3-3 機率的主要定理	52
§3-4 機率分布	58
§3-5 極限定理	70
§3-6 二元分布	78

第四章 頻率計算

§4-1 總 述	88
§4-2 選擇方法	93
§4-3 經驗頻率	95
§4-4 皮爾遜曲綫族	109
§4-5 皮爾遜第三型曲綫	118
§4-6 皮爾遜第一型曲綫	146
§4-7 正態曲綫	157

§4-8	皮尔逊第五型曲线	166
§4-9	皮尔逊第六型曲线	171
§4-10	格拉姆-夏里埃曲线	174
§4-11	勃罗夫柯维奇曲线	188
§4-12	对数正态曲线	197
§4-13	克里茨基与閔凯里曲线	205
§4-14	耿贝尔曲线	213
§4-15	古德力区曲线	225
§4-16	机率格纸	231
§4-17	频率曲线的绘制方法	244
§4-18	组合频率	253

第五章 相关分析

§5-1	总述	264
§5-2	直线相关(上)	268
§5-3	直线相关(下)	277
§5-4	曲线选配	282

第六章 抽样误差

§6-1	总述	286
§6-2	抽样误差的公式	289
§6-3	样本参数的无偏估计量	314

附录 1	Γ 函数与 β 函数	320
附录 2	换元公式的导演	324
附录 3	$\Gamma(n)$ 值的表	326
附录 4	$\log \Gamma(n)$ 值的表	331
附录 5	正态曲线下的面积及纵坐标表(一)	335
附录 6	正态曲线下的面积及纵坐标表(二)	340
附录 7	正态机率函数的二阶、三阶及四阶导函数表	345
附录 8	经验频率表 $P = \frac{m}{n+1} \times 100\%$	350
附录 9	皮尔逊第三型曲线的离均系数 ϕ 表 $0 < C_s < 3$	352
附录 10	皮尔逊第三型曲线的离均系数 ϕ 表 $C_s > 3$	355
附录 11	皮尔逊第三型曲线的纵坐标 K_P 值	356
附录 12	克里茨基与閔凯里曲线的纵坐标 K_P 值	359
附录 13	参考文献	366
附录 14	中外人、地译名对照表	373

序 言

近年来，水文計算中运用数理統計法来分析水文現象日益增多，許多工程人員也迫切要求深入了解这方面的知識。但我国目前还缺少比較完整的水文統計書籍，为此，編者收集了現在常用的各項水文統計方法，編成本書。書中除闡述各家的理論和計算的方法外，並提出了一些个人的看法。

利用数理統計法的优点是可以給出水文現象的偶然性規律。大家知道，当前的气象科学，还不能对降雨作出长期定量的預报，因此也就不能确切推算某水利工程在运用期間内会必然遭遇的各种水文特征，但是利用数理統計法我們却有可能估計这些工程在运用期中所可能遭遇的各种水文情况，解决了水利工程规划中所提出的問題。

不过，由于数理統計法的广泛应用，也使若干工程人員对这种方法产生了一些誤解。他們过高地評價了这个方法，以至形式地和机械地利用了这个方法而不去認真和客觀地分析原始資料和現象的本質，忽視了自然地理的影响，显然，这是不对的。应当指出，数理統計法的基本職能只是各种技术性的計算，其本身也帶有若干缺点，故在統計分析的同时，必須密切結合現象的物理成因分析和考虑地理分佈的綜合，这样才能产生最好的效果，否則就容易脱离实际，甚至歪曲事实。

、本書共分六章，篇首簡述了統計研究的內容，其次分別詳細地敘述了水文計算中的統計参数、机率理論、頻率計算、相关分析和抽样誤差等，并对每个細節和公式都作了詳細的推演与考証。

这里还需說明一下，即在頻率計算中，对于一般的方法，如特大值处理、抽样誤差的估計和适綫示例等，仅列于皮尔逊第三型曲綫的一节中。其他綫型的处理方法大都是相同的，只要掌握

了这些方法的来源，就不难加以类推，故書中不再贅述。又数理統計的計算方法甚为簡單，因而本書对同类的問題，也不再重复举例，以节省篇幅；有些方法仅用文字上說明已能代替例子的，也都省去了。

1956年6月間，編者写成了本書的前身“水文学中的数理統計法”，并由治淮委员会勘测設計院加以油印。去年来北京水利科学研究院工作，又接受了編写本書的任务。本書初稿完成于去年9月份，現在又作了少許的修改与补充。

初稿完成前后，中国科学院数学研究所张里千先生給我們系統地介紹了机率論与数理統計方面的知識，使我們澄清了許多机率論与数理統計上的問題，对編写本書起了較大的作用，書中某些地方并引用了张先生的講課筆記〔1—20〕。

編者自1955年初开始接触水文統計工作以来，蒙淮委及水利科学研究院領導和同志們的鼓励，以及华东水利学院和数学研究所許多同志的帮助与启发，这是完成本書的重要因素，在此表示衷心的感謝。

本書主要是根据〔1—1〕〔2—1〕〔4—4〕〔4—5〕和水利科学研究所收集的各有关頻率曲綫的文献編成的，可以作为水文工作者深入研究之用，也可以作为大专学校的水文专业学生及其他数理統計工作者的参考。

限于我的能力与水平，其中的缺点和錯誤一定很多，恳切地盼望着同志們的批評与指正。

金光炎 写于水利科学研究院

1958.2.北京

目 錄

第一章 总 論

§1-1 統計的概念	5
§1-2 統計觀察	6
§1-3 統計歸納	9
§1-4 水文計算与数理統計	12

第二章 統計参数

§2-1 总 述	13
§2-2 平均数	14
§2-3 算术平均数	16
§2-4 中位数	21
§2-5 众 数	23
§2-6 几何平均数	25
§2-7 調和平均数	26
§2-8 矩	27
§2-9 离散度	31
§2-10 偏 度	36
§2-11 峯 度	38
§2-12 簡化計算与近似公式	40

第三章 机率論的基本知識

§3-1 总 述	46
§3-2 机率的概念	47
§3-3 机率的主要定理	52
§3-4 机率分布	58
§3-5 极限定理	70
§3-6 二元分布	78

第四章 頻率計算

§4-1 总 述	88
§4-2 选样方法	93
§4-3 經驗頻率	95
§4-4 皮尔遜曲綫族	109
§4-5 皮尔遜第三型曲綫	118
§4-6 皮尔遜第一型曲綫	146
§4-7 正态曲綫	157

§4-8 皮尔逊第五型曲线	166
§4-9 皮尔逊第六型曲线	171
§4-10 格拉姆-夏里埃曲线	174
§4-11 勃罗夫柯维奇曲线	188
§4-12 对数正态曲线	197
§4-13 克里茨基与閔凯里曲线	205
§4-14 耿贝尔曲线	213
§4-15 古德力区曲线	225
§4-16 机率格纸	231
§4-17 频率曲线的绘制方法	244
§4-18 组合频率	253

第五章 相关分析

§5-1 总述	264
§5-2 直线相关(上)	268
§5-3 直线相关(下)	277
§5-4 曲线选配	282

第六章 抽样误差

§6-1 总述	286
§6-2 抽样误差的公式	289
§6-3 样本参数的无偏估计量	314

附录 1 Γ 函数与 β 函数	320
附录 2 换元公式的导演	324
附录 3 $\Gamma(n)$ 值的表	326
附录 4 $\log \Gamma(n)$ 值的表	331
附录 5 正态曲线下的面积及纵坐标表(一)	335
附录 6 正态曲线下的面积及纵坐标表(二)	340
附录 7 正态机率函数的二阶、三阶及四阶导函数表	345
附录 8 经验频率表 $P = \frac{m}{n+1} \times 100\%$	350
附录 9 皮尔逊第三型曲线的离均系数 ϕ 表 $0 < C_s < 3$	352
附录 10 皮尔逊第三型曲线的离均系数 ϕ 表 $C_s > 3$	355
附录 11 皮尔逊第三型曲线的纵坐标 K_P 值	356
附录 12 克里茨基与閔凯里曲线的纵坐标 K_P 值	359
附录 13 参考文献	366
附录 14 中外人、地译名对照表	373

第一章 总 論

§1-1 統計的概念

(一)統計一語的意義

統計一語系由拉丁文中 *Status* 導來，它原來的含義是“情況”。最初出現這個名詞是在十八世紀初期，那時，它所表達的意義十分狹窄，只是說明一些國家的政治情況，直到該世紀末葉以後，才開始流傳起來。而到現今，可以說已發展到相當豐富和十分實用的地步，並且已以一門完善的單元出現於科學的領域中。

目前統計一語已有了廣泛的含義，通常把說明社會及自然現象的全部數字叫做統計，這些數字可以反映出社會中或自然界中各種大量現象的發展水平和發展過程，以及這些現象的質量內容變化等等。

統計的內容可分為：資料的搜集，所得資料的統計研究，統計觀察方法的擬定和統計資料的分析，而後者實為數理統計學的內容。

(二)總體和樣本

總體是客觀存在的，是在同基礎上結合個別單位的集團，也就是指有一個共同的聯繫結合起來的許多因素（觀測單位）的群。而樣本是總體中由若干單位組成的小群。因此總體可以分割成很多樣本。

總體必須要包括所研究現象的全部情況，例如蚌埠站年降水量總體，必須要包括蚌埠站自古迄今再展延至未來極其長遠的歲月中所有的年降水量，而現在所掌握的有限資料，不過是總體中的一個很小部分，也就是一個樣本。

總體的概念，我們還可以理解得更廣一些，這需要決定於我們所研究的任務與內容。例如我們要研究淮河流域的年雨量，則全流域各地區的年雨量就構成一個總體；又如我們僅需研究淮河

中游汛期的降雨，那么淮河中游全部地区的汛期降雨就形成一个总体。由于研究的任务不同，故上例中所包括的地区大小、时间长短都不同。不过，以后我們所要談論的总体，都是以時間过程来表示的，也就是研究水文現象在时序上的变化。

水文現象的总体有着与别的統計現象不同的特点，即总体是无限的。例如工业成品的总体，毕竟可以知道它有多少个或者知悉一个十分近似的数目，而水文系列究竟有多长呢？除非考古到地球开始的那一天和推测到地球存在的最后一天，而現在所能获得的資料，都是无限总体中的一个很小样本。这样說来，水文系列的总体項数似乎是无法捉摸的，幸而在机率論和数理統計学中給出了估計总体的定理与方法，才使得有可能在样本数足够大时去推测总体的近似情况。因此，总体与样本的概念有着严格的区别，两者不能混淆。总体只有通过样本去認識它，而不能直接算得，尤其不能把样本認作是总体，仅在样本具有大量觀測資料时，我們才可以应用数理統計中的方法对总体作出充分精确的估計，以滿足实际分析的需要。

§1-2 統計觀察

(一)統計研究的三个阶段

所有完整的統計研究都必須包含三个工作阶段：統計觀察，統計归納和統計分析。

統計觀察是任何統計研究的开始阶段，是科学地有組織地采集和登記統計研究綱要中所需要的材料，也就是把組成总体的各个单位連同它們的所有特征进行有計劃的登記，此时要对大量現象和过程作出綜合說明。

統計归納是把由統計觀察中所得到的大量的原始資料列入綜合說明全部总体的归納表中，或繪列于統計图表內。

統計分析是把整理好的材料，进行科学的加工和分析，得出結果。而且还必須对各种現象作出总结性的描述，以及从質和量上作出多方面的分析。隨着分析就要計算各种簡單的或者复杂的

統計特征值，本書是着重介紹這個方面的知識。

統計研究的三個階段是依次連著的，相互之間都有密切的關係，實際工作時，不應忽視其中任何一個階段。

在上述的統計研究三個階段之前，應當加上統計研究的設計（研究計劃的編制）和準備工作。設計和準備工作做得好，對於統計研究工作的順利開展有着重大的意義。

（二）統計觀察的重要意義

統計觀察是統計工作的第一個階段，它的基本任務就是要從所研究總體的各個單位中獲得真實而可靠的原始資料，它的基本方法是大量觀察法。

大家知道，一切的統計研究都是在各項原始資料的基礎上建立起來的，只有取得能反映現實情況的資料，才能對所研究現象的實際研究發生作用。就統計資料的構成內容來說，每一個單位，每一個數據，都是構成資料整體的一員，唯有對全部個體的調查記錄都做到正確無誤，才能保證統計整體不發生偏差。因此，統計觀察是整個統計工作的基礎，只有先把基礎打好，才能保證整個統計工作的順利開展。如果調查工作做得不好，得到的資料不準確或殘缺不全，則不管以後的歸納和進一步的研究進行得怎樣好，而全部統計研究必遭失敗，因為原始資料的欠缺是不能在進一步的研究中消去的。統計觀察的重要意義就在於此，它是決定着統計研究成敗的關鍵。

（三）統計觀察的方法

（1）經常性觀察和一次性觀察

經常性觀察就是隨着事實的產生而隨時進行登記，例如：水文與氣象的觀測。一次性觀察就是間隔一定時間或根據需要臨時舉行的觀察，例如歷史洪水的調查，河道規劃前的水系踏勘及洪水或內澇災害後損失調查等等。

（2）全面觀察和非全面觀察

全面觀察是在調查過程中要把被調查總體的所有單位都包括進去，但不是無條件地包括一切，只是在具體任務的範圍內進行。

例如人口普查及国民經济执行情况等。在水利规划中，有时也要进行小范围的全面观察，例如，在水庫的淹沒范围内，必須对所需迁移的村莊、人口、財產等全部調查清楚。

非全面观察就是对被調查总体中的一部分单位进行观察。非全面观察应用很广，它只需要較少的人力与物力，並可应用更完善的观察法，且获得資料迅速而及时。非全面观察可分为抽样观察法、重点法和专题論述法。抽样观察是抽取总体的一部分进行观察，再把这种抽样样本的观察結果推論于总体。重点法是对总体中佔絕大比重的那些单位进行观察，而把比重小的部分予以省略。专题論述法是深入研究一个或少数几个典型的对象，虽則这种研究結果不能据以推論至总体，但可更深入、更詳細地研究大規模調查中所不能說明的問題。水文分析中以抽样观察法为最常用，我們将在第六章中单独論述。

(3) 資料来源和取得資料的方式

一、自填法：是按統計表报的形式，从机关中系統地取得必要的材料，如水文測驗成果列表汇报。

二、派人調查：有些調查工作，不能只靠机关自己填报，而需要組織一批人力去当地調查。如具有技术性的河道調查，堤壩的汛前檢查等等。

三、通訊观察：这种方法是把統計表发給被調查人，要他填写后寄回。一般实际收回来的統計表少于发出去的数目，而答案也不一定确切。这种方法，只有在个别条件下才能应用，如对資料的准确性要求不高或沒有更妥善的办法时。

(四) 統計观察的差誤及其防止方法

由于真正科学地組織全部統計工作而使材料严格地合乎客觀，乃是社会主义統計固有的特点，只有在这种条件下，統計才能發揮巨大的作用。

斯大林同志在联共(布)第十四次党代表大会上的报告中，揭露人民公敌歪曲統計材料的企图时说：“我們認為：中央統計局应当供給客觀的完全沒有任何偏見的材料，因为企图使数字去迁

就任何偏見都是犯罪行为”(斯大林全集第七卷第329頁至330頁)。因此統計材料的严格精确性乃是社会主义統計的法律。在統計工作各个阶段，特別是在調查阶段，材料都必須是严格的和客观的。

如果統計調查是通过各机关提交統計表报的，則此时調查本身的差誤来源是机关内部原始记录的錯誤，預防这种錯誤的方法是要很好地組織原始记录，經常檢查登記材料。但統計表报也可能由于故意歪曲现实，这是有意欺騙国家的行为，要受到人民輿論的指責，如其上級机关能細心檢查統計表报，便能暴露这种錯誤。

在进行統計調查过程中，由于未能正确反映事实而产生的登記上的錯誤，乃是錯誤的来源，登記的錯誤有兩类：一是故意的錯誤，如当事人有意識地歪曲事实。另一类是无意的，其来源可能是登記者偶然写錯，也可能是被調查者对問題理解不够明确或对过去某些事实記憶不清，这类錯誤多是偶然性的，在此情况下，只要能很好的組織原始材料和經常細心的檢查，虽則或許仍存有偶然性的錯誤，但对处理大量材料来源，就不甚要紧，因为在归納时，正負錯誤可能相互抵銷，結果不会相差很大，所以只要認真处理原始材料和經常細心檢查，錯誤就一定能够防止。

§1-3 統計归納

(一)統計归納与分組的意义

在經過統計观察取得了大量的原始材料以后，繼之而来的任务是将这些材料进行彙总並加以系統化。彙总的方法首先就是按照一定的标志，把登記材料分門別类地分成若干部分或集团，也就是按照某一标志划分为各种不同的組，这种分組的方法就叫做統計分組法。

分組的目的是把同一現象、同一类型加以合併，使它們与其他現象、其他类型区别开来。它的作用就是按不同的标志，把資料的特征反映出来，以便进一步运用各种統計方法进行加工，来

研究它們的規律性和相互关系。

由于資料的分組是正确地反映出社会或自然現象特征的基礎，因此对現象的各种形态必須进行具体研究，以及正确地 and 客觀地定出各种不同的組。如果分組分得不科学不正确，不管已取得的資料是如何广泛、正确和全面，我們仍不能根据这些材料得出客觀的結論，所以統計归納阶段也具有相当重大的意义。

水文計算中的分組比較簡單，只需把現象的不同类型、不同性質、不同地区(或地点)及不同作用条件区分开来，也就是說分組時必須把上述非同位相的資料分开，而把同位相的資料归納在一起。例如：降水量与逕流量是屬於不同类型，春汛流量与秋汛流量的性質是不同的；人类活动影响前后的水文資料，由于产生的条件不同，因此都不能把它們归納在一起。

为了計算方便，同位相資料也要按数值的不同分成若干組，但是由于水文資料的項数都很少，不能用分組來計算，故很少应用。但在某些情况下，資料經過綜合而系列加长了，則仍能应用，如対站年法的計算和超定量法的分析等。

(二)統計表

原始材料經過統計归納后，其結果需要用数字來說明，所以我們可將整理和分組的結果，列于縱橫表格內，以便使統計特征一目了然，这种表式就称为統計表。

用統計表，可有很多优点：

(1)避免了文字上重复解釋的麻煩，故用統計表常較长篇大論为生动。

(2)錯誤及遺漏的地方容易发现，便于及时檢查。

(3)統計参数的計算和分析工作可以簡化，並利于数字間的比較。

(4)表中將資料按次序排列，有明显的邏輯系統，为作統計分析的依据，又其所表現的內容、規律和相互关系，易得明确的概念和容易記憶。

(三)統計图