

实用气象统计方法-笔记

一、绪论

1. 数学模型（模式、表达式）

- (1) 统计学模式（经验模式）
- (2) 数值模拟模式（动力学模式、理论模式）
- (3) 动力-统计模式（半经验半理论模式）

2. 气象统计方法

(1) 统计分析（重要）：指对大气科学研究对象的概率、概率分布及其特征量进行统计，或进行相关分析、类型分析、随机过程和随机场分析以及试验设计等。

(2) 统计推断：对由样本资料得到的关于总体的概率、概率分布和相应的定量关系进行估计、判断，看它们是否能真正代表总体，包括估计和假设检验两个方面。

3. 气象统计方法的基本任务

1. 初步加工和整理有关数据（订正、数据中心化、标准化、重要特征量的统计（平均值、标准差）的统计）
2. 定量分析气象变化规律（研究对象随时间、空间的变化规律，建立相应的经验（统计）模型）
3. 判断由有限数据统计得出的某种规律或关系是否符合实际情况

二、气象数据的初步整理与加工

1. 气象数据的审查（准确性、均一性、代表性、比较性）

一般来自气象台站网观测和整编的气象数据都是经过审查的，对于来源于其他方面的气象数据，使用前是需要认真对待的。

2. “野点”处理与气象数据数列订正

- (1) 识别野点：假设检验→①舍去②订正或更正
- (2) 序列订正方法（相邻站点）

①差值订正法②比值订正法③维尔达（Wild）综合订正法④回归订正法

- (3) 超短序列订正方法（相邻站点）

①全概率公式②简单的线性内插③谐波分析

3. 气象数据的特征量统计

- (1) 气象数据的集中性和离散性
- (2) 频率、频数分布

频率分布直方图

- (3) 累积频率（常用于代替或估计概率分布函数）

累积频率曲线图

4. 常用的气象要素数据特征量统计

- (1) 描述气象数据集中位置的特征量

①算术平均值 \bar{x} ②众数 M_0 ③中位数 M_e

- (2) 描述气象数据离散程度的特征量

①距平

②平均绝对变率 V_d （平均差、平均误差）

距平绝对值的平均值，表示整个观测数据对于平均值 \bar{x} 的离散程度

③平均相对变率 V_r

平均绝对变率与平均值的百分比，一般来说平均相对变率比平均绝对变率可以更好地反映数据的离散程度

④极差（较差、全距或振幅）

某气象要素最大值 X_{max} 与最小值 X_{min} 之差

⑤均方差（标准差） S

⑥离差系数 C_v （变差系数、变异系数）

均方差与样本平均值之比

（3）描述气象数据分布对称性的特征量

①偏度系数②峰度系数

5.气象数据的初步加工

（1）中心化（零均值化）→化为距平值的形式，消除平均值的影响（平均值为0）

（2）标准化→消除了平均值和量纲的影响（平均值为0，标准差为1）

极值归一化处理：把原来的数据都化为0~1的数值

（3）等级化（编码法）

（4）滑动平均

可以起到滤波的作用，消除一部分随机因素的干扰

（5）差分

可以使整个随机过程达到平稳化

（6）正态化

使数据成为正态分布的形式

6.有效数字及其舍入规则

准确度： x 位有效数字

精确度：具有 y 位小数的精确度（精确到 y 位小数）

三、气象数据的统计推断

1.气象数据特征量的点估计和区间估计

（1）点估计：对于任何一个待估计的总体参数，往往可有若干个估计量，从中确定一个较好的估计量近似看作总体参数，称为参数的点估计或定值估计。

例：使用样本平均值估计一个服从正态分布随机变量总体的数学期望

较好的估计量所具有的性质：

①一致性②无偏性③有效性④充分性

（2）区间估计：构造一个置信区间，使被估计参数以一定的概率 $P=1-\alpha$ 落在此区间内

2.假设检验：检验总体假设的具体方法

假设检验的一般步骤：

①根据问题的要求做出假设 H_0 ；

②设计（或选用）满足上述两个条件的统计量 Ω ；

③给定一个信度 α ，根据 Ω 已知分布函数查出置信限，确定临界区间 R_α ，使在假设 H_0 成立时有 $P(\Omega \in R_\alpha) = \alpha$ ；

④根据样本资料计算出 Ω 的具体值 ω ；

⑤判断：如 $\omega \in R_\alpha$ 在临界区间内（置信区间外），则拒绝原假设 H_0 ；反之，则接受原假设 H_0 。

3.常用的气象统计假设检验方法：

（1） u 检验

适用于统计量 Ω 服从 $N(\mu, \sigma)$ 正态分布（高斯分布）的各种假设检验

1) 数学期望等于某指定值时的检验（假设 $\mu=\mu_0$ ）

给定信度 α ，查正态分布双侧置信限表可得置信限 u_α ，使其满足 $P(|u| \geq u_\alpha) = \alpha$ ，若 $|u| < u_\alpha$ 则接受原假设

$$u = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

2) 概率等于某指定值的检验 (假设 $p = p_0$)，对研究对象进行了 n 次观测求得频率 v

$$u = \frac{v - p_0}{\sqrt{\frac{p_0(1 - p_0)}{N}}}$$

给定信度 α ，查正态分布双侧置信限表可得置信限 u_α ，使其满足 $P(|u| \geq u_\alpha) = \alpha$ ，若 $|u| < u_\alpha$ 则接受原假设

3) 两个数学期望相等的检验

4) 两个概率相等时检验

(2) t 检验

适用于统计量 Ω 服从 t 分布 (与 u 分布类似) 的各种假设检验，通常在小样本的情况下采用 (10 次左右)

1) 数学期望等于某个值时的检验

2) 两个数学期望相等时的检验

3) 线性相关系数的检验

(3) 卡方检验

适用于统计量 Ω 服从 $N(\mu, \sigma)$ 卡方分布 (χ^2 分布) 的各种假设检验

1) 总体方差等于某指定值时的检验

2) 总体方差小于等于某指定值时的检验

3) 总体方差大于等于某指定值时的检验

4) 列联表独立性检验

5) 总体分布的假设检验

*检验研究对象 X 服从什么类型的分布：皮尔逊 (Pearson) 定理

(4) F 检验

适用于统计量 Ω 服从 $N(\mu, \sigma)$ F 分布的各种假设检验

1) 两个总体方差相等的检验

2) 多个数学期望相等的检验

(5) 符号检验

适用于所要检验的对象观测数据成对出现的情形，但不要求它们一定服从正态分布

(6) 秩和检验

主要适用于两个总体分布是否相同的检验 (考虑观测值有次序的检验方法)、

四、气象科学试验结果的方差分析

方差分析方法的应用 (所检验的各个总体都应服从正态分布)：

① 分析研究对象在不同试验条件或处理下，观测结果之间的差异。判断这些差异是状态效应还是随机效应造成的。从而了解哪些条件或因子是重要的、关键的。

② 分析研究对象 (气象要素、条件或现象等) 是否存在周期性变化。

五、气象领域中的相关分析方法

1. 相关系数：

简单相关系数、偏相关系数

2. 回归分析：

一元线性回归分析、多元线性回归分析、逐步回归分析

六、气象领域中的趋势分析

1. 线性倾向
2. 滑动平均
3. 累积距平

七、气象场的经验正交分解（EOF 分析）

把一个三维场分解为空间模态和时间序列, 方差贡献率越高的模态为该地区该气象场的主要空间分布形式。其高值区域反映该区域该气象要素变化大。