



关于海水温度基本参数

μ 与 σ 的点线面估计

孙琪田

(大连760试验场)

海水温度是海水的重要物理因子之一,观测、分析和研究它的变化规律,无疑是海洋探索中的重要一环。本文仅基于数理统计方法,对海水温度的基本参数——均值 μ 与标准偏差 σ (或方差 σ^2)进行统计估计。由于海水温度 T 是个服从正态变化的随机量^[1],所以文中有关 μ 和 σ 的点、线、面的估计,均以高斯分布理论为基本依据。

一、 μ 与 σ 的点估计

假定我们对温度随机变量 T 进行了 n 次独立观测,分别得到

$$t_1, t_2, \dots, t_n.$$

如果 θ 是温度总体分布的某个参数,只要我们能找到一子样函数 $\hat{\theta}$ 来估计 θ ,则 $\hat{\theta}$ 就称为 θ 的点估计。具体说由于 $T \in N(\mu, \sigma^2)$,而参数 μ 和 σ 又未知,现要从总体抽出的这一子样 $t_1, t_2, \dots, t_n$ 中来求取 μ 、 σ 的点估计量 $\hat{\mu}$ 、 $\hat{\sigma}$,其数学表达式即为:

$$\hat{\mu} = \frac{1}{n}(t_1 + t_2 + \dots + t_n) = \frac{1}{n} \sum_{i=1}^n t_i = \bar{t}, \quad (1)$$

$$S^2 = \frac{1}{n} \sum_{i=1}^n (t_i - \bar{t})^2 = \frac{1}{n} \left[\sum_{i=1}^n t_i^2 - n\bar{t}^2 \right]. \quad (2)$$

可以证明^[2](1)式是 μ 的无偏估计量,而 σ^2 的无偏估计量则为

$$\hat{\sigma}^2 = \frac{n}{n-1} S^2 = \frac{1}{n-1} \sum_{i=1}^n (t_i - \bar{t})^2 \quad (3)$$

由于点估计量 $\hat{\mu}$ 与 $\hat{\sigma}$ 也是一个随机量,所以为了描述不同取样所得估计点,我们提出如

下的线估计(区间估计)问题。

二、 μ 与 σ 的线估计

既然 $\hat{\mu}$ 和 $\hat{\sigma}$ 为随机量,它必然随观测容量数而变。 $\hat{\mu}$ 和 $\hat{\sigma}$ 在某一给定置信率中,能够包括总体参数 μ 和 σ 的区间长度究竟有多长?我们姑且把估算这一变化长度称之为线估计。计算线估计长度的基本公式为

$$P(\theta_1 < \theta < \theta_2) = 1 - \alpha. \quad (4)$$

式中 θ 仍如前所述,代表温度总体中的某个参数(如 μ , σ , σ^2 等), θ_1 和 θ_2 表示在给定置信度 α 情况下, θ 取值的上、下限, $1 - \alpha$ 为置信率,因此,对于 μ 与 σ 便有:

$$P\left(\left|\frac{\bar{t} - \mu}{\sigma/\sqrt{n}}\right| < u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha, \quad (5)$$

$$P\left[\frac{\hat{\sigma}^2(n-1)}{\lambda_2} < \sigma^2 < \frac{\hat{\sigma}^2(n-1)}{\lambda_1}\right] = 1 - \alpha. \quad (6)$$

其置信区间便为:

$$\left(\bar{t} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{t} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right), \quad (7)$$

$$\left(\hat{\sigma} \sqrt{\frac{n-1}{\lambda_2}}, \hat{\sigma} \sqrt{\frac{n-1}{\lambda_1}}\right). \quad (8)$$

通常(7)式中的 σ 是未知的,所以采用下式^[5]来计算 μ 的置信区间,即:

$$\left(\bar{t} - u_{\nu, 1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n-1}}, \bar{t} + u_{\nu, 1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n-1}}\right). \quad (9)$$

上述式中 $\hat{\sigma}$ 为子样标准差,由(3)式求得,它遵从自由度为 $n-1$ 的 χ^2 分布。 $u_{1-\frac{\alpha}{2}}$ 和

$u_{\nu, 1-\frac{\alpha}{2}}$ 可分别根据 α 查正态分布表和由给定

置信率及自由度 ν 查“学生氏” t 分布表⁽⁵⁾求得。 λ_1 和 λ_2 是在给定 $1-\alpha$ 时所选取 σ 置信区间长度较短的置信限,由给定置信率和自由度查 χ^2 分布表⁽⁵⁾获得,也可由文献(3)所介绍的公式来求得。

现分别令 μ 和 σ 的置信上、下限为:

$$\mu_{\text{上}} = \bar{x} - u_{\nu, 1-\frac{\alpha}{2}} \hat{\sigma} / \sqrt{n-1}, \quad (10)$$

$$\mu_{\text{下}} = \bar{x} + u_{\nu, 1-\frac{\alpha}{2}} \hat{\sigma} / \sqrt{n-1}. \quad (11)$$

$$\sigma_{\text{上}} = \hat{\sigma} \sqrt{\frac{n-1}{\lambda_2}}, \quad (12)$$

$$\sigma_{\text{下}} = \hat{\sigma} \sqrt{\frac{n-1}{\lambda_1}}. \quad (13)$$

则 μ 和 σ 的线估计长度即为:

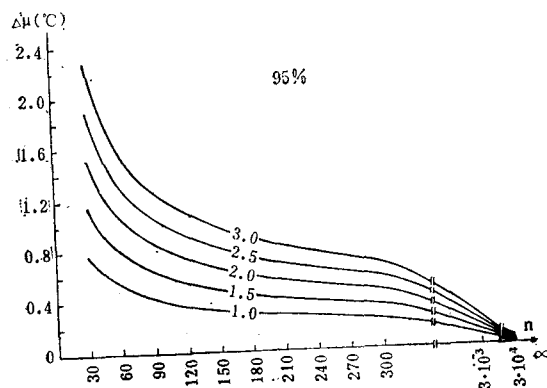


图 1

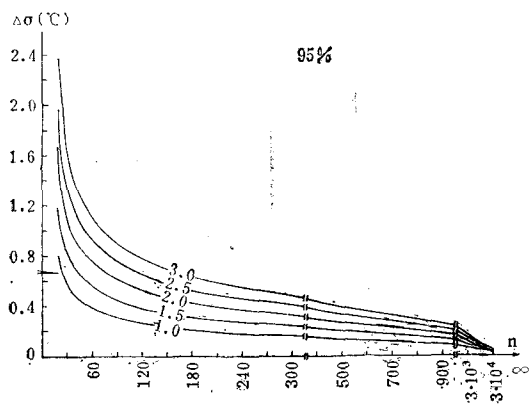


图 2

$$\Delta \mu = \mu_{\text{下}} - \mu_{\text{上}} = 2u_{\nu, 1-\frac{\alpha}{2}} \hat{\sigma} / \sqrt{n-1}, \quad (14)$$

$$\Delta \sigma = \sigma_{\text{下}} - \sigma_{\text{上}} = \hat{\sigma} \sqrt{n-1} \left(\frac{1}{\sqrt{\lambda_1}} - \frac{1}{\sqrt{\lambda_2}} \right) \quad (15)$$

由(14)和(15)式分别绘制了置信率为95%时的 $\Delta \mu$ 与 n 和 $\Delta \sigma$ 与 n (在 $\hat{\sigma}$ 为1.0, 1.5, 2.0, 2.5, 3.0五种情况下)的关系图(图1和图2),以及 $\Delta \mu$ 与 $\hat{\sigma}$ 跟 $\Delta \sigma$ 与 $\hat{\sigma}$ (在不同 n 值的情况下)的关系图(图3和图4)。

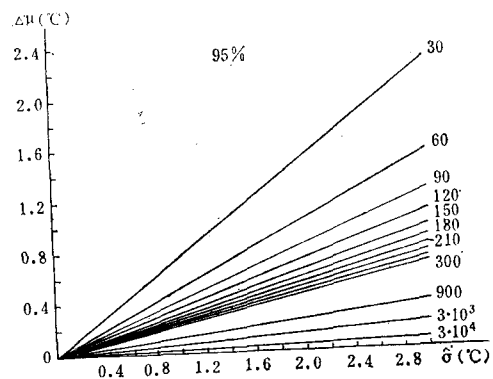


图 3

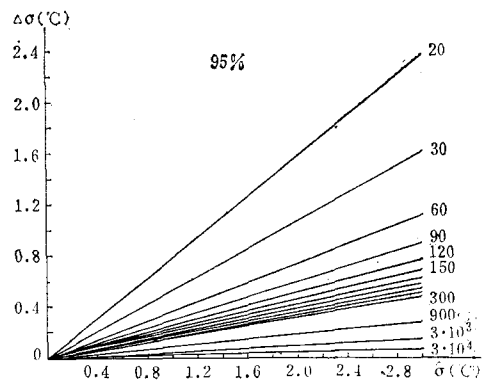


图 4

三、 μ 和 σ 联合置信面的估计

根据上面已经给出的 μ 和 σ 的置信区间估算式(5)和(6),可以写成:

$$P\left(-u_{1-\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} < u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha,$$

$$P\left(\lambda_1 < \frac{\hat{\sigma}^2(n-1)}{\sigma^2} < \lambda_2 = 1 - \alpha.\right)$$

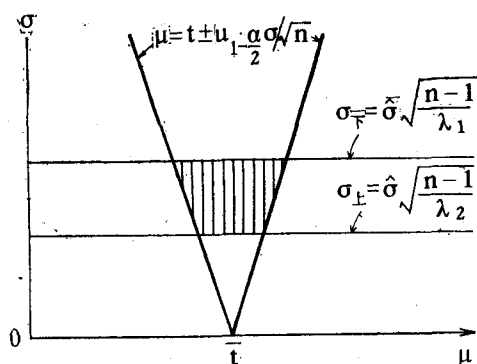


图 5

我们知道变量 $\bar{t}$ 和 $\bar{\sigma}$ 的分布是独立分布，则 μ 和 σ 联合置信率的分布，即为⁽⁴⁾：

$$P \left[-u_{1-\frac{\alpha}{2}} < \frac{\bar{t}-\mu}{\sigma/\sqrt{n}} < u_{1-\frac{\alpha}{2}}, \right. \\ \left. \lambda_1 < \frac{\hat{\sigma}^2(n-1)}{\sigma^2} < \lambda_2 \right] = (1-\alpha)^2. \quad (16)$$

那么 μ 与 σ 的联合置信面，即如图 5（阴影部分）所示。若令阴影的面积为 ΔQ ，其计算公式可根据梯形原理推出：

$$\Delta Q = 2u_{1-\frac{\alpha}{2}} \cdot \hat{\sigma} \sqrt{\frac{n-1}{n}} \left(\frac{1}{\sqrt{\lambda_1}} + \frac{1}{\sqrt{\lambda_2}} \right) \cdot \hat{\sigma} \sqrt{n-1} \left(\frac{1}{\sqrt{\lambda_1}} - \frac{1}{\sqrt{\lambda_2}} \right) \cdot \frac{1}{2} = u_{1-\frac{\alpha}{2}}.$$

$$\cdot \frac{\hat{\sigma}^2(n-1)}{\sqrt{n}} \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2} \right). \quad (17)$$

此处的 ΔQ 勿须给出它的具体量纲。它的实际含意可理解在取样均值为 $\bar{t}$ 及取样标准差为 $\hat{\sigma}$ 时，

$$\left. \begin{aligned} \sigma & \text{ 从 } \hat{\sigma} \sqrt{\frac{n-1}{\lambda_2}} \rightarrow \hat{\sigma} \sqrt{\frac{n-1}{\lambda_1}} \\ \mu & \text{ 从 } \bar{t} \mp u_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{\lambda_2}} \sqrt{\frac{n-1}{n}} \\ & \rightarrow \bar{t} \mp u_{1-\frac{\alpha}{2}} \cdot \frac{\hat{\sigma}}{\sqrt{\lambda_1}} \sqrt{\frac{n-1}{n}} \end{aligned} \right\} \text{ 置信率}$$

为 $1-\alpha$ 条件下，其联合置信率为 $(1-\alpha)^2$ 时， μ 与 σ 的联合置信域的面估计量。

现以北黄海—海洋水文观测站的冬春夏秋代表月（2、5、8、11）表层水温为例，对 μ 和 σ 于不同年份（不同观测数）的点、线、面的估计值列入下表。

四、结 语

通过对海水温度主要参数 μ 与 σ 的点、线、面的估计可知，估计量的准确程度与取样数 (n) 有关，当 $n \rightarrow \infty$ 时，点估计量 $\hat{\mu}$ 与 $\hat{\sigma}$ 就趋向真值 μ 与 σ ，而围绕真值 μ 与 σ 的线估计长度 $\Delta\mu$ 和 $\Delta\sigma$ 与面估计值 ΔQ 都趋向于零，这时在以 σ 为纵轴、 μ 为横轴的坐标系中，对某一特定海区、特定时间上的温度的点、线、面的

μ 和 σ 的点、线、面估计值表

月 份		2				5				8				11			
年 份		63	62	62	62	63	62	62	62	63	62	62	62	63	62	62	62
观 测 数 (n)			65	71	77		65	71	77		65	71	77		65	71	77
		28	113	283	452	31	124	310	496	31	124	310	496	30	120	300	480
点估计值 (°C)	$\hat{\mu}$	1.1	1.8	1.1	1.1	9.7	10.2	9.8	9.7	23.9	23.0	22.6	22.3	12.0	12.1	12.4	12.3
	$\hat{\sigma}$	0.53	1.04	1.10	1.07	1.44	1.64	1.67	1.58	0.55	1.11	1.22	1.45	1.86	1.88	1.73	1.75
线估计值 (°C)	$\Delta\mu$	0.42	0.39	0.26	0.20	1.07	0.58	0.37	0.28	0.41	0.40	0.27	0.26	1.41	0.68	0.39	0.31
	$\Delta\sigma$	0.30	0.28	0.18	0.14	0.79	0.42	0.27	0.20	0.30	0.28	0.19	0.18	1.02	0.49	0.28	0.22
面估计值 ΔQ		0.13	0.11	0.05	0.03	0.86	0.25	0.10	0.06	0.13	0.11	0.05	0.05	1.46	0.33	0.11	0.07

估计值仅表现为一点。这就是说,在当 $n \rightarrow \infty$ 时,仅作点估计值就可以了。

然而,在实际问题中,由于我们对海水温度的观测取样总是有限的,要想无穷尽的取样那是根本办不到的,这就要求人们如何选择最佳观测数,以达到通过子样了解总体的目的。

总之,我们认为对海水温度基本参数 μ 与 σ 作点线面的定量估计,不仅可帮助我们有效地选取最佳观测数,以获取在给定条件下令人满意的统计结果,而且对海水温度的中、长期预报方面,也有着一定的参考价值。

参 考 文 献

- [1] 水文组, 1981。试验海区表层水温的统计析。海上试验与测量技术1(2): 1—3。
- [2] 王梓坤, 1979。概率论基础及其应用。科学出版社, 219页, 223—225页。
- [3] 孙琪田等, 1982。北隍城、老虎滩、小长山地区海水表层温度的某些统计特性。海洋通报1(4): 8。
- [4] A. M. 穆德等(史定华译), 1978。统计学导论。科学出版社, 289页。
- [5] Bethea, R.M., 1975. Statistical Methods for Engineers and Scientists, Marcel Dekker, Inc. New York and Basel. p. 135, pp. 518—532.