

2070 Temperature Prediction

吴千里

更新于 06 December, 2021

目录

1	假设 Assumption	2
2	温度数据	3
3	预测区间	5
4	(忽略) 仅 Deterministic 的预测	6
5	包含 Stochasticity 的预测	8

1 假设 Assumption

1. 0W-12W（经度），54N-70N（纬度）看作能影响鱼群的海域面积
2. 不考虑未来出现的意外/人为因素

2 温度数据

```
# 设置需要的经纬度范围 (可更改)
# Set the bounds of regions here
longitude_lowerbound <- 0
longitude_upperbound <- 12
latitude_lowerbound <- 54
latitude_upperbound <- 70

# Get the average temperature for each year
data <- data %>% group_by(Time) %>% summarise(avg_temp = mean(Sea.Surface.Temperature))
# Print data
data

## # A tibble: 35 x 2
##   Time avg_temp
##   <dbl>   <dbl>
## 1  1981     6.69
## 2  1982     7.82
## 3  1983     7.62
## 4  1984     8.07
## 5  1985     7.95
## 6  1986     7.86
## 7  1987     8.14
## 8  1988     8.27
## 9  1989     8.08
## 10 1990     8.33
## # ... with 25 more rows
```

```
plot(data$Time,  
      data$avg_temp,  
      col = "blue",  
      main = "Sea Surface Temperature of 0-12W and 54-70N from 1981-2015",  
      xlab = "Year",  
      ylab = "Temperature (celsius)")
```

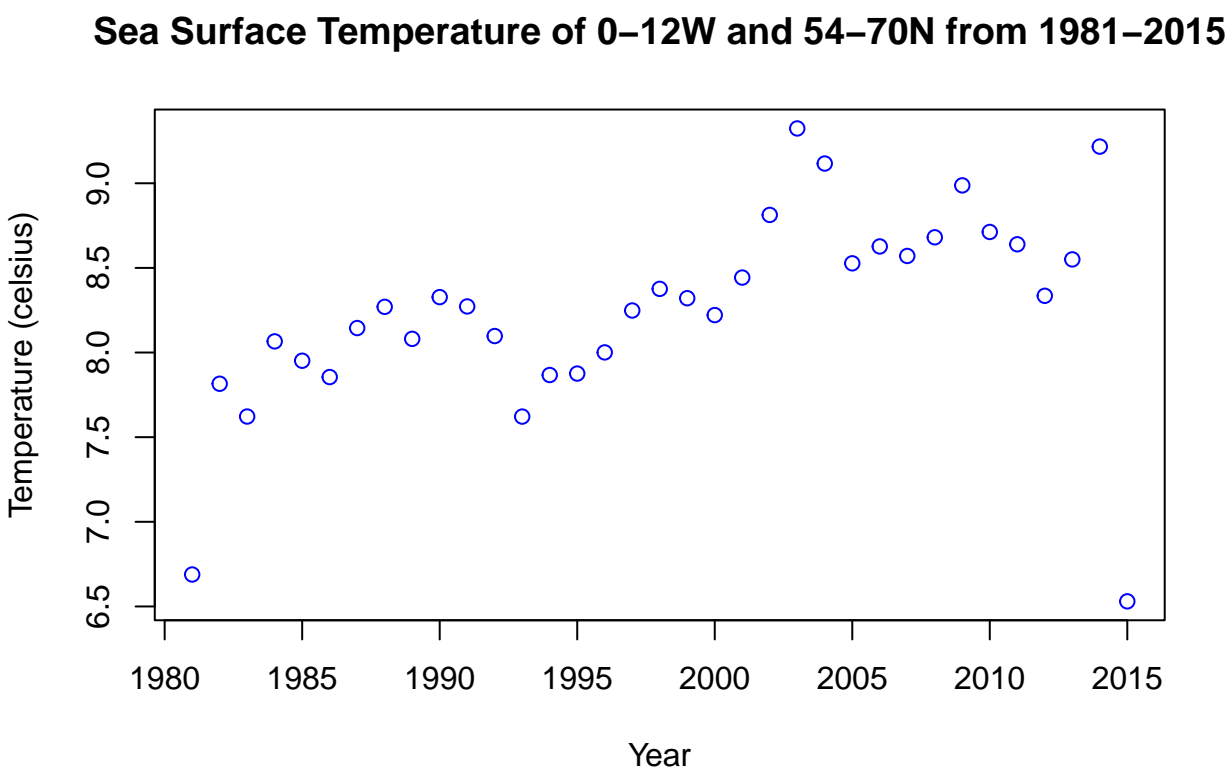


图 1: 年平均海表温度 (0-12W & 54-70N)

3 预测区间

A prediction interval is a type of confidence interval (CI) used with predictions in regression analysis; it is a range of values that predicts the value of a new observation, based on your existing model.

The formula of prediction interval:

$$\hat{y}_h \pm t_{\frac{\alpha}{2}, n-2} \cdot \sqrt{MSE(1 + \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum (s_i - \bar{x})^2})}$$

```
predict(models, newdata = new.year, interval = "prediction")
```

```
##      fit      lwr      upr
## 1 10.2526  8.57538 11.92981
```

Our 95% Prediction Interval is (8.58, 11.93), which means according to our regression model, the average sea surface temperature in 2070 will fall into the range of 8.58°C to 11.93°C 95% of time.

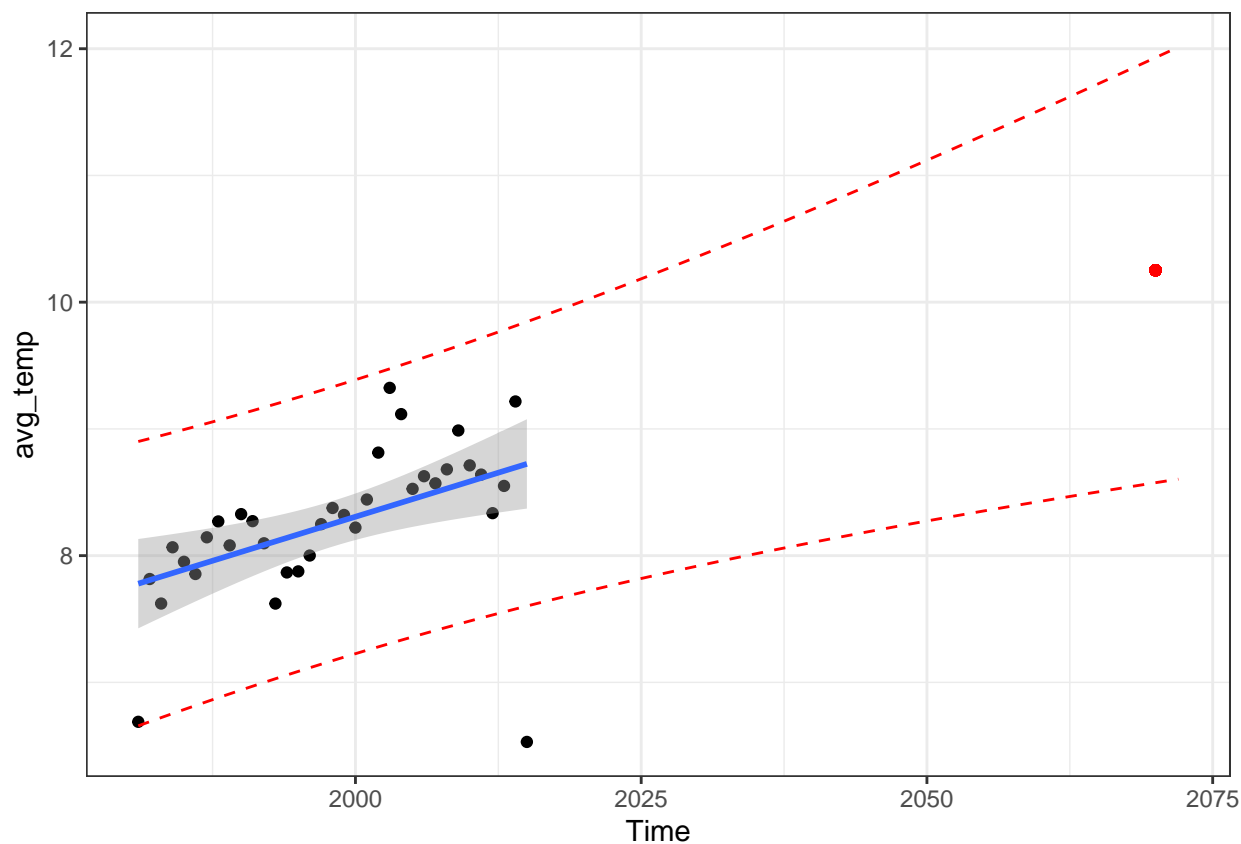


图 2: 2070 年的预测区间以及预期数值（红点）

4 (忽略) 仅 Deterministic 的预测

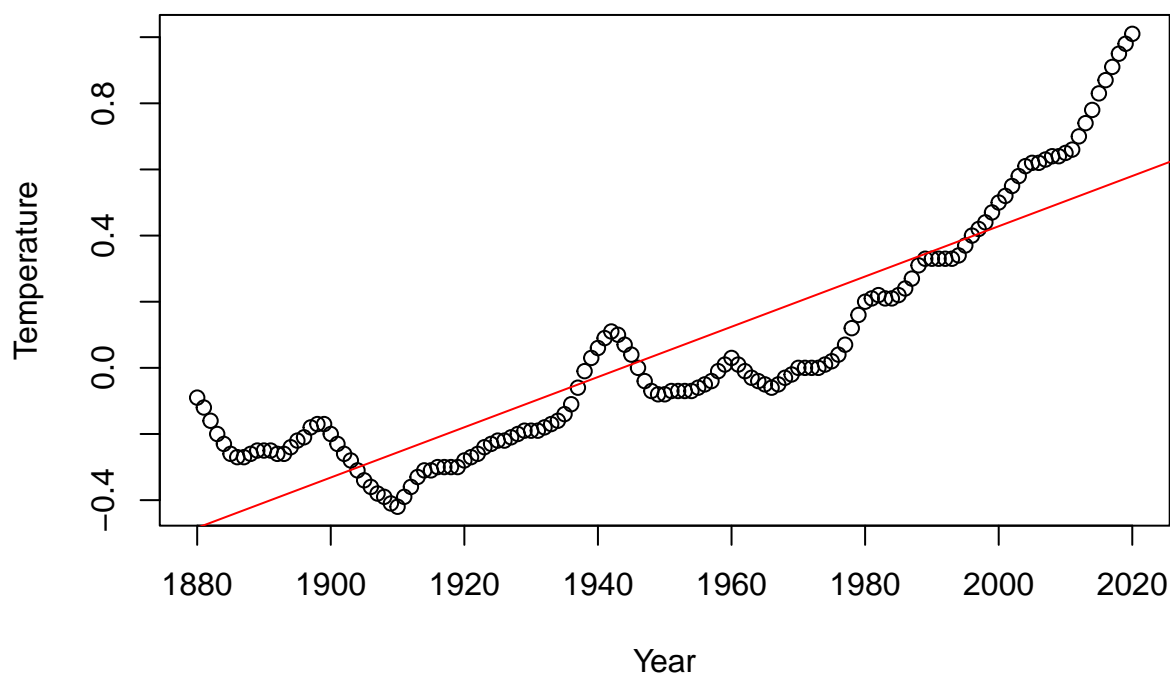


图 3: 1880-2020 海表温度变化趋势

上图应使用非线性的拟合曲线，或者可以仅使用近 50 年的数据，如下页面所示：

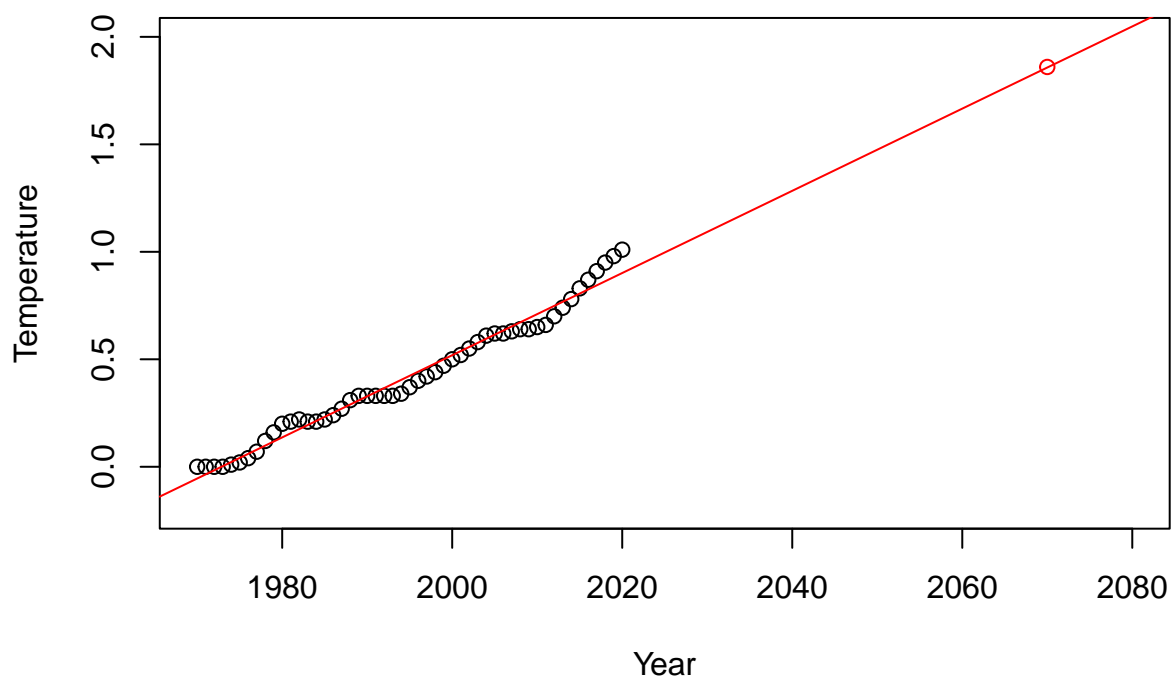


图 4: 1970-2020 年海表温度数据拟合曲线

上图的**红点**代表 2070 年拟合曲线预测的海水温度，大约为 **+1.86 °C**

5 包含 Stochasticity 的预测

5.0.1 方法一：通过 R 语言自带公式得出预测区间

缺点：不清楚其统计原理

优点：简单明了，且数据比较合理

```
temperature <- allTemp$Temperature[70:141]
year <- allTemp$Year[70:141]
models <- lm(temperature ~ year)
models

##
## Call:
## lm(formula = temperature ~ year)
##
## Coefficients:
## (Intercept)          year
##   -29.15034         0.01483

new.year <- data.frame(
  year = c(2070)
)
predict(models, newdata = new.year, interval = "prediction")

##          fit          lwr          upr
## 1 1.556521 1.368443 1.7446
```

总结：预期值是 1.56°C，预测区间是 1.37°C - 1.74°C。

5.0.2 方法二：看作 discrete model

每年温度变化量是一个取自 1950-2020 年温度增加量的随机数，模拟到 2070 年

缺点： 不确定变成 discrete 是否合理 + 温度偏高

优点： 过程清晰 + 和作业题相似

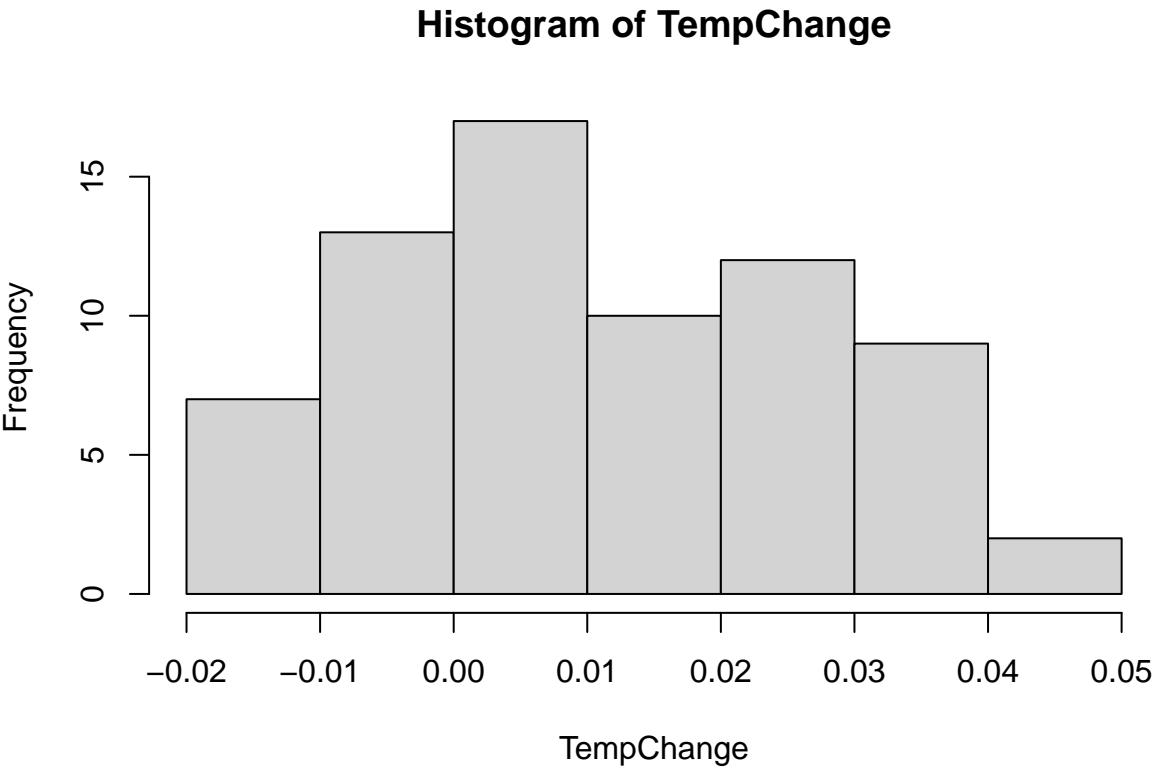


图 5: 每年温度变化量的频率分布

```
# 1950-2020 年间每年温度变化量（后一年温度减前一年）
TempChange

## [1] 0.01 0.00 0.00 0.00 0.01 0.01 0.01 0.03 0.02 0.02 -0.02 -0.02
## [13] -0.02 -0.01 -0.01 -0.01 0.01 0.02 0.01 0.02 0.00 0.00 0.00 0.01
## [25] 0.01 0.02 0.03 0.05 0.04 0.04 0.01 0.01 -0.01 0.00 0.01 0.02
## [37] 0.03 0.04 0.02 0.00 0.00 0.00 0.00 0.01 0.03 0.03 0.02 0.02
## [49] 0.03 0.03 0.02 0.03 0.03 0.03 0.01 0.00 0.01 0.01 0.00 0.01
## [61] 0.01 0.04 0.04 0.04 0.05 0.04 0.04 0.04 0.03 0.03

# 上述数据的总结：
summary(TempChange)

##      Min.   1st Qu.   Median     Mean  3rd Qu.     Max.
## -0.02000  0.00000  0.01000  0.01557  0.03000  0.05000
```

```
# 上述数据的方差:
```

```
sd(TempChange)
```

```
## [1] 0.01724804
```

5.0.3 predict temperature in 2070 using discrete model

```
# 模拟 50 次的分布
```

```
tempPrediction(x = 50)
```

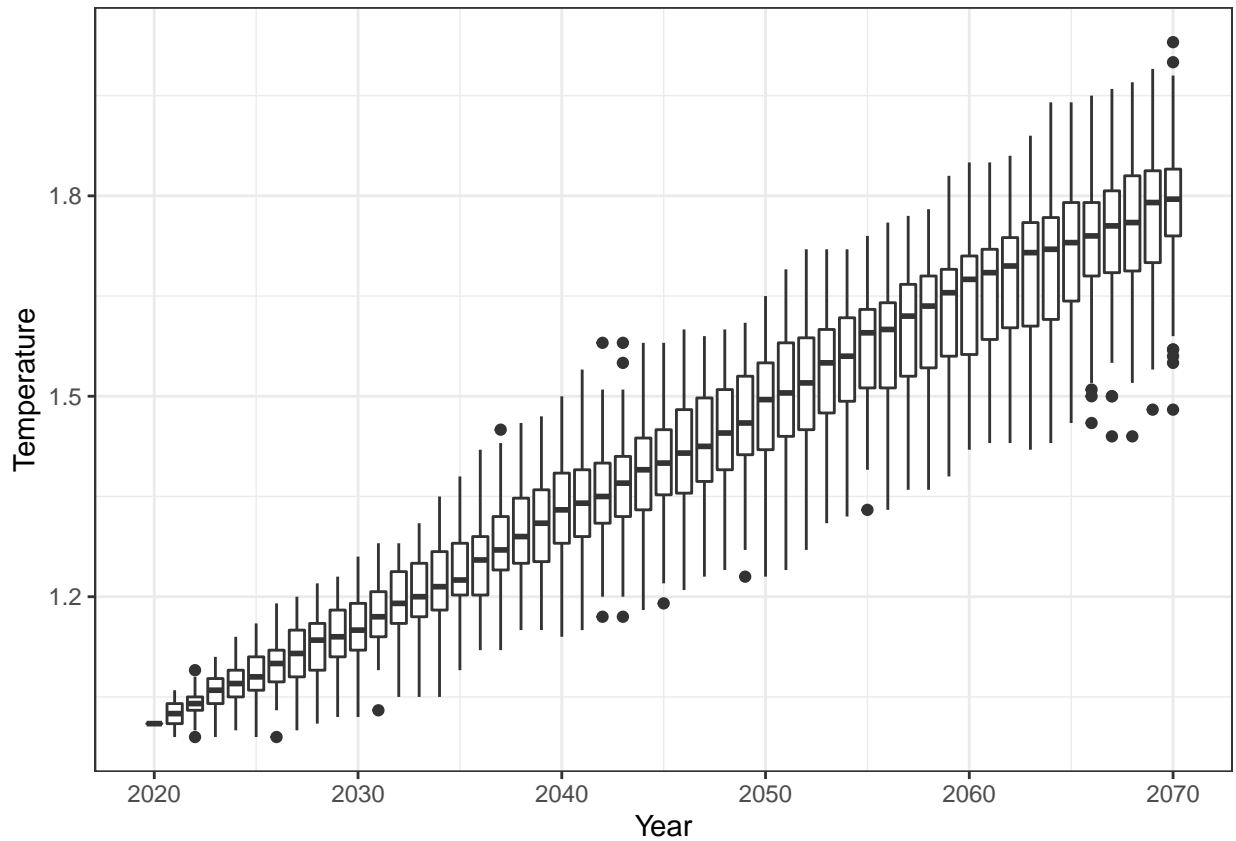


图 6: 50 次模拟 2070 年温度分布

总结: 2070 年海水表面温度预期值是 1.79°C, 预测区间为 1.7°C - 1.9°C 之间。