

习题 7.5

1. 在检验了一个车间生产的 20 个轴承外座圈的内径后得到下面数据 (单位: mm)

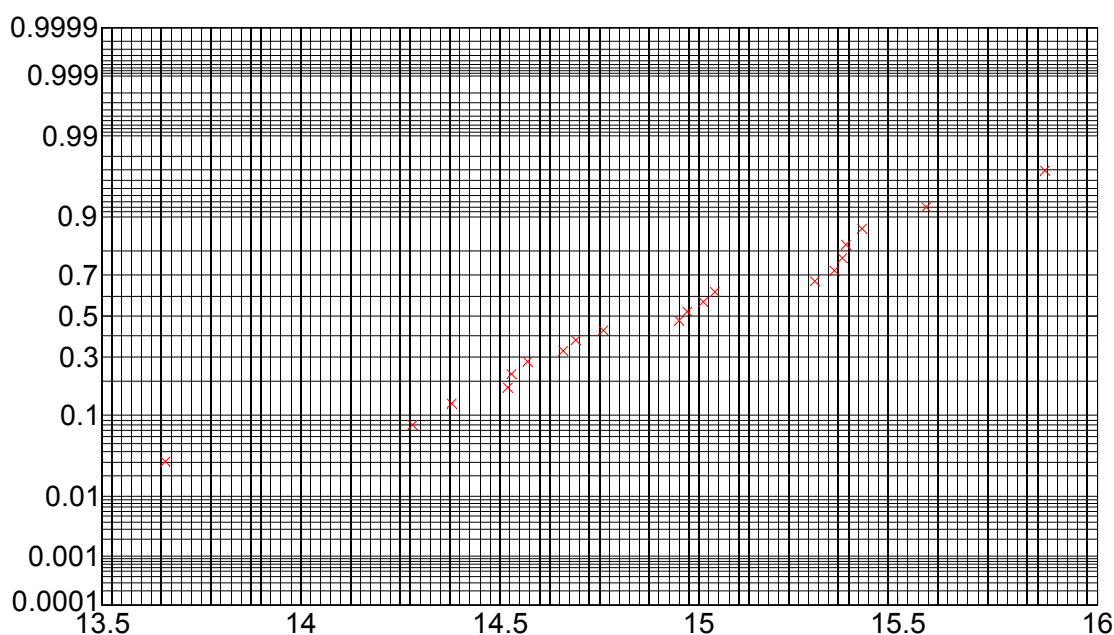
15.04 15.36 14.57 14.53 15.57 14.69 15.37 14.66 14.52 15.41
15.34 14.28 15.01 14.76 14.38 15.87 13.66 14.97 15.29 14.95

(1) 作正态概率图, 并作初步判断;

(2) 请用 W 检验方法检验这组数据是否来自正态分布 ($\alpha=0.05$) ?

解: (1) 将数据按从小到大的顺序排列, 并计算修正频率 $\frac{i-3/8}{n+1/4}$, $i=1, 2, \dots, n$, 且 $n=20$,

数据	13.66	14.28	14.38	14.52	14.53	14.57	14.66	14.69	14.76	14.95
修正频率	0.0309	0.0802	0.1296	0.1790	0.2284	0.2778	0.3272	0.3765	0.4259	0.4753
数据	14.97	15.01	15.04	15.29	15.34	15.36	15.37	15.41	15.57	15.87
修正频率	0.5247	0.5741	0.6235	0.6728	0.7222	0.7716	0.8210	0.8704	0.9198	0.9691



所描点近似在一条直线上, 初步判断这组数据来自正态分布总体;

(2) 假设 H_0 : 数据来自正态分布总体,

$$\text{选取统计量 } W = \frac{\left[\sum_{i=1}^n (a_i - \bar{a})(X_{(i)} - \bar{X}) \right]^2}{\sum_{i=1}^n (a_i - \bar{a})^2 \sum_{i=1}^n (X_{(i)} - \bar{X})^2} = \frac{\left[\sum_{i=1}^{[n/2]} a_i (X_{(n+1-i)} - X_{(i)}) \right]^2}{\sum_{i=1}^n (X_{(i)} - \bar{X})^2},$$

显著性水平 $\alpha=0.05$, $W_{\alpha}(n) = W_{0.05}(20) = 0.905$, 左侧拒绝域 $W = \{w \leq 0.905\}$,

将数据按从小到大的顺序排列, 并列出 W 检验的系数 $a_i(20)$,

数据	13.66	14.28	14.38	14.52	14.53	14.57	14.66	14.69	14.76	14.95
$a_i(20)$	0.4734	0.3211	0.2565	0.2085	0.1686	0.1334	0.1013	0.0711	0.0422	0.0140
数据	14.97	15.01	15.04	15.29	15.34	15.36	15.37	15.41	15.57	15.87
$a_i(20)$	-0.0140	-0.0422	-0.0711	-0.1013	-0.1334	-0.1686	-0.2085	-0.2565	-0.3211	-0.4734

有 $\bar{x}=14.9115$, 计算可得 $w=0.9743 \notin W$,

故接受 H_0 , 拒绝 H_1 , 即可以认为这组数据来自正态分布总体.

2. 抽查克矽平治疗矽肺患者 10 名, 得到他们治疗前后的血红蛋白量之差如下:

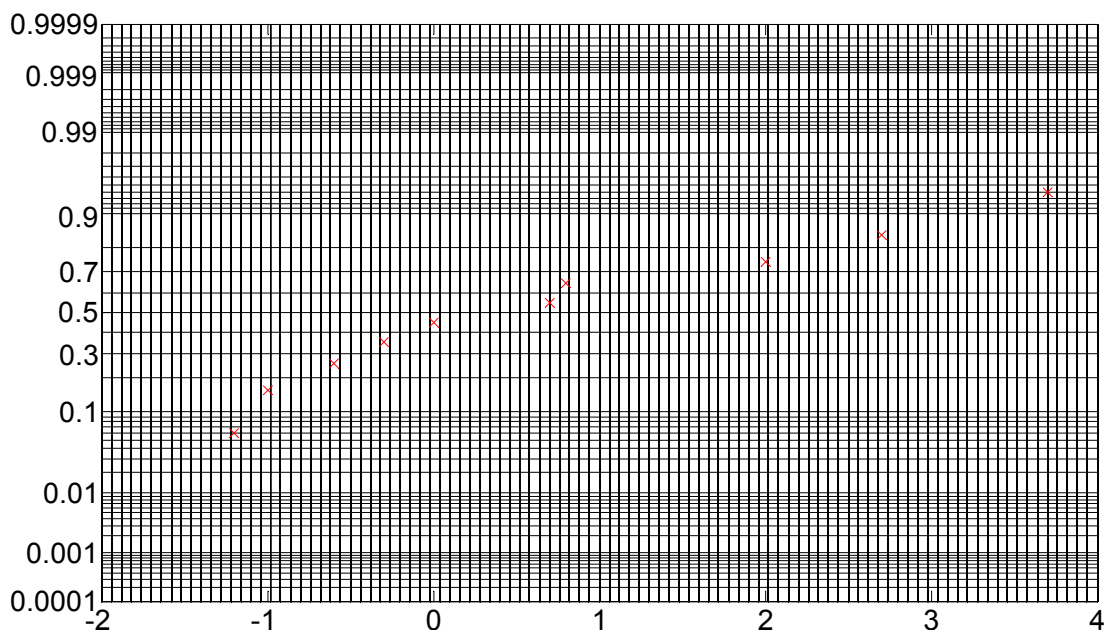
2.7 -1.2 -1.0 0 0.7 2.0 3.7 -0.6 0.8 -0.3

(1) 作正态概率图, 并作初步判断;

(2) 请用 W 检验方法检验治疗前后的血红蛋白量之差是否来自正态分布 ($\alpha = 0.05$) ?

解: (1) 将数据按从小到大的顺序排列, 并计算修正频率 $\frac{i-3/8}{n+1/4}$, $i=1, 2, \dots, n$, 且 $n=10$,

数据	-1.2	-1.0	-0.6	-0.3	0	0.7	0.8	2.0	2.7	3.7
修正频率	0.0610	0.1585	0.2561	0.3537	0.4512	0.5488	0.6463	0.7439	0.8415	0.9390



所描点近似在一条直线上, 初步判断这组数据来自正态分布总体;

(2) 假设 H_0 : 数据来自正态分布总体,

$$\text{选取统计量 } W = \frac{\left[\sum_{i=1}^n (a_i - \bar{a})(X_{(i)} - \bar{X}) \right]^2}{\sum_{i=1}^n (a_i - \bar{a})^2 \sum_{i=1}^n (X_{(i)} - \bar{X})^2} = \frac{\left[\sum_{i=1}^{[n/2]} a_i (X_{(n+1-i)} - X_{(i)}) \right]^2}{\sum_{i=1}^n (X_{(i)} - \bar{X})^2},$$

显著性水平 $\alpha = 0.05$, $W_{\alpha}(n) = W_{0.05}(10) = 0.842$, 左侧拒绝域 $W = \{w \leq 0.842\}$,

将数据按从小到大的顺序排列, 并列出 W 检验的系数 $a_i(10)$,

数据	-1.2	-1.0	-0.6	-0.3	0	0.7	0.8	2.0	2.7	3.7
$a_i(10)$	0.5739	0.3291	0.2141	0.1224	0.0399	-0.0399	-0.1224	-0.2141	-0.3291	-0.5739

有 $\bar{x} = 0.68$, 计算可得 $w = 0.9252 \notin W$,

故接受 H_0 , 拒绝 H_1 , 即可以认为这组数据来自正态分布总体.

3. 某种岩石中的一种元素的含量在 25 个样本中为

0.32 0.25 0.29 0.25 0.28 0.30 0.23 0.23 0.40 0.32 0.35 0.19 0.34
0.33 0.33 0.28 0.28 0.22 0.30 0.24 0.35 0.24 0.30 0.23 0.22

有人认为该样本来自对数正态分布总体, 请用 W 检验方法作检验 ($\alpha = 0.05$).

解: 设总体 X 服从对数正态分布 $LN(\mu, \sigma^2)$,

则 $Y = \ln X$ 服从正态分布 $N(\mu, \sigma^2)$,

假设 H_0 : 数据来自正态分布总体,

$$\text{选取统计量 } W = \frac{\left[\sum_{i=1}^n (a_i - \bar{a})(Y_{(i)} - \bar{Y}) \right]^2}{\sum_{i=1}^n (a_i - \bar{a})^2 \sum_{i=1}^n (Y_{(i)} - \bar{Y})^2} = \frac{\left[\sum_{i=1}^{\lfloor n/2 \rfloor} a_i (Y_{(n+1-i)} - Y_{(i)}) \right]^2}{\sum_{i=1}^n (Y_{(i)} - \bar{Y})^2},$$

显著性水平 $\alpha = 0.05$, $W_{\alpha}(n) = W_{0.05}(25) = 0.918$, 左侧拒绝域 $W = \{w \leq 0.918\}$,

将数据按从小到大的顺序排列, 并列出 W 检验的系数 $a_i(25)$,

数据 Y_i	-1.6607	-1.5141	-1.5141	-1.4697	-1.4697	-1.4697	-1.4271	-1.4271	-1.3863	-1.3863
$a_i(25)$	0.4450	0.3069	0.2543	0.2148	0.1822	0.1539	0.1283	0.1046	0.0823	0.0610
数据 Y_i	-1.2730	-1.2730	-1.2730	-1.2379	-1.2040	-1.2040	-1.2040	-1.1394	-1.1394	-1.1087
$a_i(25)$	0.0403	0.0200	0	-0.0200	-0.0403	-0.0610	-0.0823	-0.1046	-0.1283	-0.1539
数据 Y_i	-1.1087	-1.0788	-1.0498	-1.0498	-0.9163					
$a_i(25)$	-0.1822	-0.2148	-0.2543	-0.3069	-0.4450					

有 $\bar{y} = -1.2794$, 计算可得 $w = 0.9687 \notin W$,

故接受 H_0 , 拒绝 H_1 , 即可以认为数据 Y_i 来自正态分布总体, 即原数据来自对数正态分布总体.

4. 对第 3 题的数据, 试用 EP 检验方法检验这些数据是否来自正态总体 (取 $\alpha = 0.05$).

解: 假设 H_0 : 数据来自正态分布总体,

$$\text{选取统计量 } T_{EP} = 1 + \frac{n}{\sqrt{3}} + \frac{2}{n} \sum_{i=2}^n \sum_{j=1}^{i-1} \exp\left\{-\frac{(x_j - x_i)^2}{2s_*^2}\right\} - \sqrt{2} \sum_{i=1}^n \exp\left\{-\frac{(x_i - \bar{x})^2}{4s_*^2}\right\},$$

显著性水平 $\alpha = 0.05$, $T_{1-\alpha, EP}(n) = T_{0.95, EP}(25) = 0.370$, 右侧拒绝域 $W = \{w \geq 0.370\}$,

计算可得 $T_{EP} = 0.0831 \notin W$,

故接受 H_0 , 拒绝 H_1 , 即可以认为这些数据来自正态分布总体.