Chap. 8 #1, 2, 3, 7, 8, 13

1. Johnson et al. (1970) considered the behavior of a cenosphere-resin composite under hydrostatic pressure. The authors pointed out that most deep submersible vehicles utilize a buoyancy material, known as *syntactic foam*, that is a composite of closely packed hollow glass microspheres embedded in a resin matrix. These microspheres are relatively expensive to manufacture, and the cost of the syntactic foam is principally determined by the cost of the microspheres. The authors also noted that the ash from generating stations burning pulverized coal contains a small proportion of hollow glassy microspheres, known as *cenospheres*, and these have about the right size distribution for use in syntactic foam. The cenospheres can be readily collected from the ash-disposal method used in certain British generating stations. The authors were thus interested in whether the cenospheres would, in some applications, perform as well as the manufactured microspheres.

In attempting to assess the usefulness of cenospheres as a component of syntactic foam, Johnson et al. investigated the effects of hydrostatic pressure (such as exists in the ocean depths) on the density of a cenosphere-resin composite. The results are given in Table 9.2. What is the P-value for a test of $H_0 : \beta = 0$ against the alternative $\beta > 0$ for these data?

Table 9.2 The Effects of Hydrostatic Pressure on the Density of a Cenosphere-Resin Composite

Specimen	Pressure (psi) X_i	Density (g/cm^3) Y_i
1	0	0.924
2	5,000	0.988
3	10,000	0.992
4	15,000	1.118
5	20,000	1.133
6	25,000	1.145
7	30,000	1.157
8	100,000	1.357

Source: A. A. Johnson, K. Mukherjee, S. Schlosser, and E. Raask (1970).

$H_0: \beta = 0$

$H_a: \beta > 0.$

Using theil regression, by R:

```
> X = c(0,5000,10000,15000,20000,25000,30000,100000)
> Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
> theil(X, Y, beta.0 = 0, type = "u")
Alternative: beta greater than 0
C = 28, C.bar = 1, P = 0
beta.hat = 0
alpha.hat = 0.975

1 - alpha = 0.95 upper bound for beta:
-Inf, 0
```

Theil stat $C = 28$, $\bar{C} = 1$

$p\text{-val} = 0$.

$\therefore$ We reject H_0 and conclude $\beta > 0$ that the effects of hydrostatic pressure on the density of a cenosphere-resin composite is positive.

2. Explain why the effect of the unknown intercept parameter α (See model (9.1)) is "eliminated" in the application of procedure (9.4) to a set of data.

Assumptions

A1. Our straight-line model is

$$Y_i = \alpha + \beta x_i + e_i, \quad i = 1, \ldots, n, \tag{9.1}$$

where the x's are known constants and α (the intercept) and β (the slope) are unknown parameters.

A2. The random variables $e_1, \ldots, e_n$ are a random sample from a continuous population that has median 0.

$$C = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c(D_j - D_i), \tag{9.4}$$

By (9.1), the straight-line model is

$$Y_i = \alpha + \beta X_i + e_i, \quad i = 1, \ldots, n$$

By (9.3), $D_i = Y_i - \beta_0 X_i$

$$\therefore D_j - D_i = Y_j - \beta_0 X_j - Y_i + \beta_0 X_i$$

$$= \alpha + \beta X_j + e_j - \beta_0 X_j - \alpha - \beta X_i - e_i + \beta_0 X_i$$

$$= (\beta - \beta_0) X_j - (\beta - \beta_0) X_i + (e_j - e_i) \quad ①$$

By (9.4), $C = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} c(D_j - D_i)$ and α is eliminated by ①

3. Consider the tapeworm data discussed in Problem 8.1. Using the mean weight of the initial force-fed cysticerci as the independent (predictor) variable, test the hypothesis that there was virtually no change in the mean weight of the cysticerci over the 20-day period following introduction into the dogs against the alternative that the typical tapeworm grew in size during the period of the study.

Table 8.3 Relation Between Weight of the Cysticerci of *Taenia hydatigena* Fed to Dogs and Weight of Worms Recovered at 20 Days

	Mean weight, mg	
Dog	Cysticerci	Worms recovered
1	28.9	1.0
2	32.8	7.7
3	12.0	7.3
4	9.9	7.9
5	15.0	1.1
6	38.0	3.5
7	12.5	18.9
8	36.5	33.9
9	8.6	28.6
10	26.8	25.0

Source: D. W. Featherston (1971).

$H_0: \beta = 0$

$H_a: \beta > 0.$

By R:

```
> Cys = c(28.9,32.8,12.0,9.9,15.0,38.0,12.5,36.5,8.6,26.8)
> Worms=c(1.0,7.7,7.3,7.9,1.1,3.5,18.9,33.9,28.6,25.0)
> theil(Cys, Worms, beta.0 = 0, type = "u")
Alternative: beta greater than 0
C = -7, C.bar = -0.156, P = 0.758
beta.hat = -0.157
alpha.hat = 11.143

1 - alpha = 0.95 upper bound for beta:
-Inf, 0.625
```

$\therefore$ p-val $= 0.758$

$\therefore$ We fail to reject H_0 and conclude that there was virtually no change in the mean weight of the cysticerci.

7. Estimate β for the cenosphere-resin data of Table 9.2.

Table 9.2 The Effects of Hydrostatic Pressure on the Density of a Cenosphere-Resin Composite

Specimen	Pressure (psi)	Density (g/cm^3)
1	0	0.924
2	5,000	0.988
3	10,000	0.992
4	15,000	1.118
5	20,000	1.133
6	25,000	1.145
7	30,000	1.157
8	100,000	1.357

Source: A. A. Johnson, K. Mukherjee, S. Schlosser, and E. Raask (1970).

By R:

```
> X = c(0,5000,10000,15000,20000,25000,30000,100000)
> Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
> theil(X, Y, slopes = T)
Alternative: beta not equal to 0
C = 28, C.bar = 1, P = 0
beta.hat = 0
alpha.hat = 0.975

All slopes:
 i j          s.ij
 1 2 1.280000e-05    1
 1 3 6.800000e-06    2
 1 4 1.293333e-05    3
 1 5 1.045000e-05    4
 1 6 8.840000e-06    5
 1 7 7.766667e-06    6
 1 8 4.330000e-06    7
 2 3 8.000000e-07    8
 2 4 1.300000e-05    9
 2 5 9.666667e-06    10
 2 6 7.850000e-06    11
 2 7 6.760000e-06    12
 2 8 3.884211e-06    13
 3 4 2.520000e-05    14
 3 5 1.410000e-05    15
 3 6 1.020000e-05    16
 3 7 8.250000e-06    17
 3 8 4.055556e-06    18
 4 5 3.000000e-06    19
 4 6 2.700000e-06    20
 4 7 2.600000e-06    21
 4 8 2.811765e-06    22
 5 6 2.400000e-06    23
 5 7 2.400000e-06    24
 5 8 2.800000e-06    25
 6 7 2.400000e-06    26
 6 8 2.826667e-06    27
 7 8 2.857143e-06    28
```

Since $N = 8(7)/2 = 28$.

$\therefore\ k = \dfrac{28}{2} = 14.$

$\therefore\ \hat{\beta} = \dfrac{S^{(14)} + S^{(15)}}{2} = 5.545 \times 10^{-6}$ by R

8. Compute the least squares estimator $\bar{\beta}$ (See Comment 5) for the cenosphere-resin data of Table 9.2, and compare $\bar{\beta}$ with the $\hat{\beta}$ value obtained in Problem 7. In general, which of $\hat{\beta}$ and $\bar{\beta}$ is easier to compute?

Table 9.2 The Effects of Hydrostatic Pressure on the Density of a Cenosphere-Resin Composite

Specimen	Pressure (psi)	Density (g/cm^3)
1	0	0.924
2	5,000	0.988
3	10,000	0.992
4	15,000	1.118
5	20,000	1.133
6	25,000	1.145
7	30,000	1.157
8	100,000	1.357

Source: A. A. Johnson, K. Mukherjee, S. Schlosser, and E. Raask (1970).

To compute the least squares estimator, by R:

```
> X = c(0,5000,10000,15000,20000,25000,30000,100000)
> Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
> summary(lm(Y~X))

Call:
lm(formula = Y ~ X)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.077517 -0.038910  0.002531  0.047584  0.057810 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.002e+00  2.730e-02  36.689 2.74e-08 ***
X           3.911e-06  6.969e-07   5.613  0.00136 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0584 on 6 degrees of freedom
Multiple R-squared:   0.84,	Adjusted R-squared:  0.8134 
F-statistic: 31.51 on 1 and 6 DF,  p-value: 0.001365
```

$\therefore \bar{\beta} = 3.911 \times 10^{-6}$

By Q7, $\hat{\beta} = 5.545 \times 10^{-6} < \bar{\beta}$.

In general, the least-square estimator is easier to calculate.

13. Obtain a 90% confidence interval for β for the cenosphere-resin data in Table 9.2.

Table 9.2 The Effects of Hydrostatic Pressure on the Density of a Cenosphere-Resin Composite

Specimen	Pressure (psi)	Density (g/cm^3)
1	0	0.924
2	5,000	0.988
3	10,000	0.992
4	15,000	1.118
5	20,000	1.133
6	25,000	1.145
7	30,000	1.157
8	100,000	1.357

Source: A. A. Johnson, K. Mukherjee, S. Schlosser, and E. Raask (1970).

By R:

```
> X = c(0,5000,10000,15000,20000,25000,30000,100000)
> Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
> confint(lm(Y~X), level = 0.9)
                    5 %         95 %
(Intercept) 9.484733e-01 1.054561e+00
X           2.557368e-06 5.265685e-06
```

$\therefore$ 90% CI for β is $(2.557 \times 10^{-6}, 5.266 \times 10^{-6})$

ph1855_hw9_ygu5

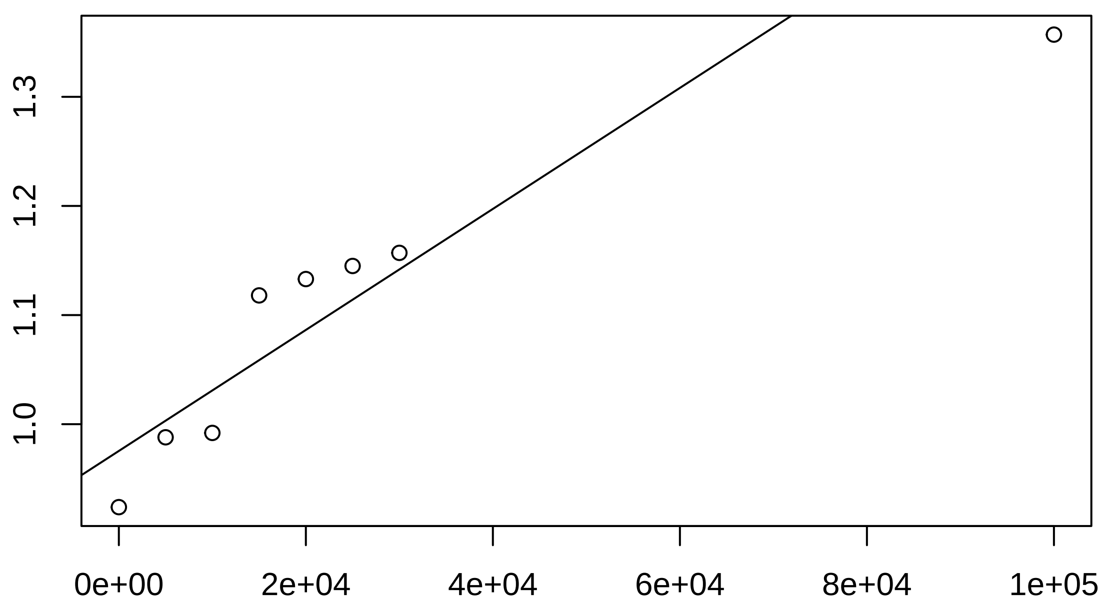
Yue Gu

2024-04-13

Hollander et al. Nonparametric Statistical Methods 2014

P457 Q1

```r
X = c(0,5000,10000,15000,20000,25000,30000,100000)
Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
theil(X, Y, beta.0 = 0, type = "u")
```

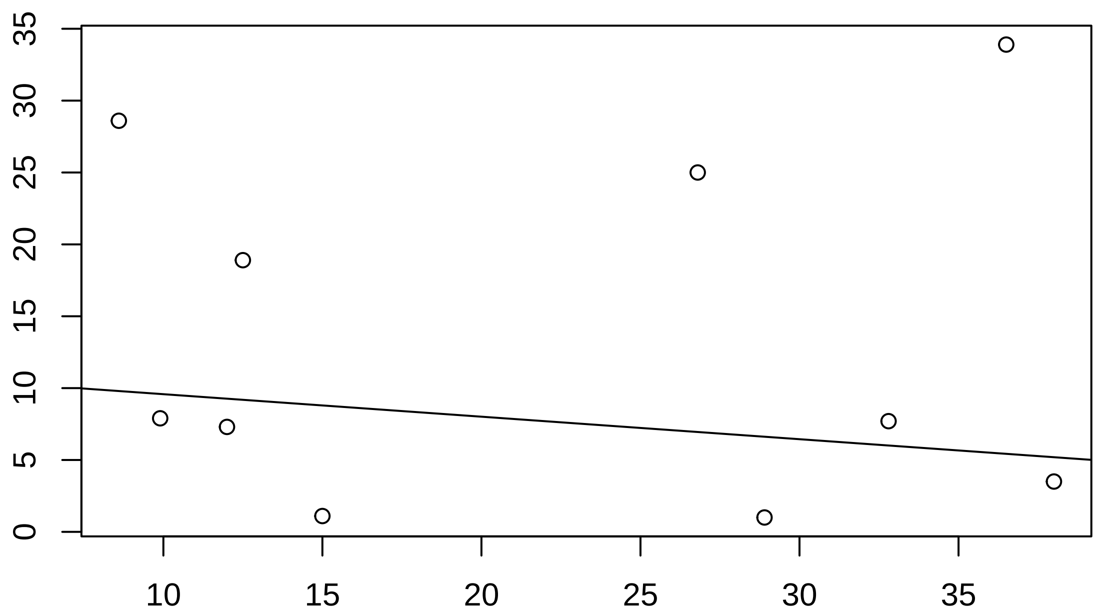



```
## Alternative: beta greater than 0
## C = 28, C.bar = 1, P = 0
```

```
## beta.hat = 0
## alpha.hat = 0.975
##
## 1 - alpha = 0.95 upper bound for beta:
## -Inf, 0
```

P457 Q3

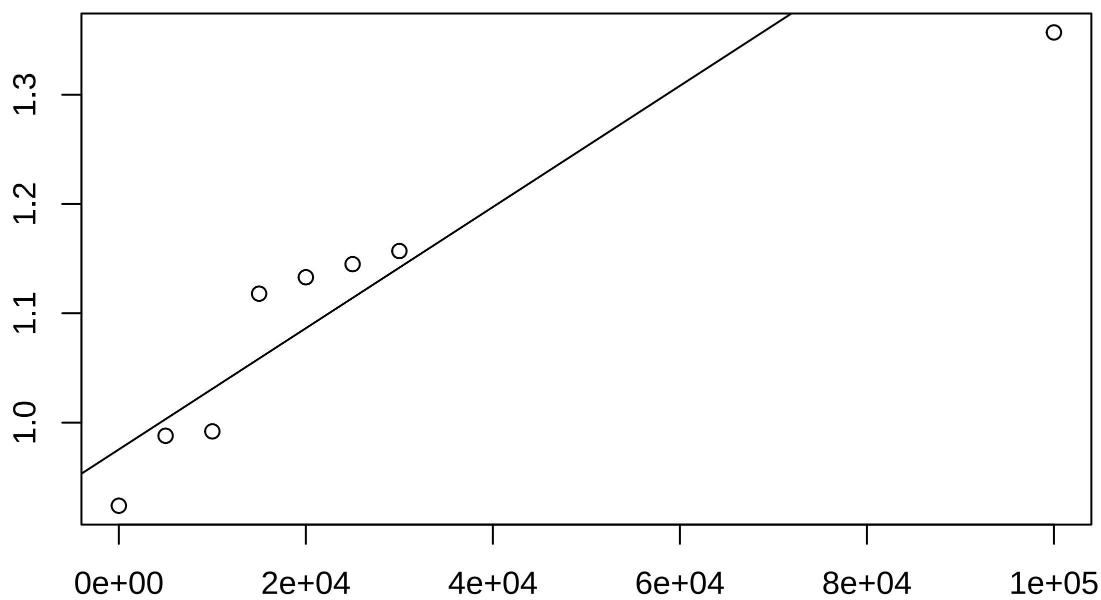
```r
Cys = c(28.9,32.8,12.0,9.9,15.0,38.0,12.5,36.5,8.6,26.8)
Worms=c(1.0,7.7,7.3,7.9,1.1,3.5,18.9,33.9,28.6,25.0)
theil(Cys, Worms, beta.0 = 0, type = "u")
```




```
## Alternative: beta greater than 0
## C = -7, C.bar = -0.156, P = 0.758
## beta.hat = -0.157
## alpha.hat = 11.143
##
## 1 - alpha = 0.95 upper bound for beta:
## -Inf, 0.625
```

P460 Q7

```r
X = c(0,5000,10000,15000,20000,25000,30000,100000)
Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
theil(X, Y, slopes = T)
```




```
## Alternative: beta not equal to 0
## C = 28, C.bar = 1, P = 0
## beta.hat = 0
## alpha.hat = 0.975
##
## All slopes:
##  i j         S.ij
##  1 2 1.280000e-05
##  1 3 6.800000e-06
##  1 4 1.293333e-05
##  1 5 1.045000e-05
##  1 6 8.840000e-06
##  1 7 7.766667e-06
##  1 8 4.330000e-06
##  2 3 8.000000e-07
##  2 4 1.300000e-05
##  2 5 9.666667e-06
##  2 6 7.850000e-06
```

```
##  2 7 6.760000e-06
##  2 8 3.884211e-06
##  3 4 2.520000e-05
##  3 5 1.410000e-05
##  3 6 1.020000e-05
##  3 7 8.250000e-06
##  3 8 4.055556e-06
##  4 5 3.000000e-06
##  4 6 2.700000e-06
##  4 7 2.600000e-06
##  4 8 2.811765e-06
##  5 6 2.400000e-06
##  5 7 2.400000e-06
##  5 8 2.800000e-06
##  6 7 2.400000e-06
##  6 8 2.826667e-06
##  7 8 2.857143e-06
##
##
## 1 - alpha = 0.95 two-sided CI for beta:
## 0, 0
```

```r
# calculate the median of the slopes
median(c(1.280000e-05,
6.800000e-06,
1.293333e-05,
1.045000e-05,
8.840000e-06,
7.766667e-06,
4.330000e-06,
8.000000e-07,
1.300000e-05,
9.666667e-06,
7.850000e-06,
6.760000e-06,
3.884211e-06,
2.520000e-05,
1.410000e-05,
1.020000e-05,
8.250000e-06,
4.055556e-06,
3.000000e-06,
2.700000e-06,
2.600000e-06,
2.811765e-06,
2.400000e-06,
2.400000e-06,
2.800000e-06,
2.400000e-06,
2.826667e-06,
2.857143e-06))
```

```
## [1] 5.545e-06
```

P460 Q13

```r
X = c(0,5000,10000,15000,20000,25000,30000,100000)
Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
summary(lm(Y~X))
```

```
##
## Call:
## lm(formula = Y ~ X)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.077517 -0.038910  0.002531  0.047584  0.057810
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.002e+00  2.730e-02  36.689 2.74e-08 ***
## X           3.911e-06  6.969e-07   5.613  0.00136 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0584 on 6 degrees of freedom
## Multiple R-squared:   0.84,  Adjusted R-squared:  0.8134
## F-statistic: 31.51 on 1 and 6 DF,  p-value: 0.001365
```

P460 Q7

```r
X = c(0,5000,10000,15000,20000,25000,30000,100000)
Y = c(0.924,0.988,0.992,1.118,1.133,1.145,1.157,1.357)
confint(lm(Y~X), level = 0.9)
```

```
##                      5 %         95 %
## (Intercept) 9.484733e-01 1.054561e+00
## X           2.557368e-06 5.265685e-06
```