

SOP 23

정도 평가에 사용되는 통계 기법

1. 대상 및 적용 분야

이 절차는 정도 평가에 쓰이는 여러 가지 통계 계산법에 대해 다룬다. 일련의 변수에 대한 평균과 표준 편차 계산, 반복 측정 자료 세트에 대한 표준 편차 계산, 평균에 대한 신뢰 구간 계산, 선택된 확률로 볼 때 평균이나 표준 편차가 유의하게 다른지를 판정하는 방법, 그리고 선형회귀의 기울기와 절편을 계산하는 방법을 상세하게 다루었다.

2. 원리

이 연산은 통계학적 원리, 특히 정규 분포에 근거한다. 해당 통계법에 대한 상세한 배경 지식은 참고문헌에 실려있다.

3. 분석 절차

3.1 측정 시리즈의 평균과 표준 편차 추정

n 개를 측정하였다면,

$$x_1, x_2, x_3, \dots, x_n$$

평균 \bar{x} 는 다음과 같이 주어진다.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

그리고 표준 편차 s 는 다음과 같이 추정된다.

$$s = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \right)^{1/2} \quad (2)$$

3.2 반복 측정한 다른 자료 세트에 대한 표준 편차 추정 s

반복 측정을 k 번 했다면,

$$d_1, d_2, d_3, \dots, d_k$$

표준편차의 추정값 s 는 다음과 같이 주어진다.

$$s_R = \left(\frac{\sum_{i=1}^k d_i^2}{2k} \right)^{1/2} \quad (3)$$

이것은 단기 표준 편차 또는 측정의 재현성¹을 잰 것이다.

3.3 평균 신뢰 구간

사용되는 공식은

$$\bar{x} \pm \frac{ts}{\sqrt{n}} \quad (4)$$

식에서

\bar{x} = 시료의 평균값,

n = 평균을 구하는데 쓰인 측정 횟수,

s = 표준 편차 추정값²,

t = Student t 값, 즉, 희망 신뢰 구간과 s 의 자유도에 상응한 확률 요인
(값은 이 절차 별첨의 표 4-9 참조),

3.4 두 평균 값의 비교

경우 1. 표준 편차가 다를 이유가 없는 경우.

¹ 국제표준기구(ISO)는 정밀도를 두 가지로 기술하고 있다: (1) **재현성(reproducibility)**, 같은 방법으로 다른 조건 (예, 다른 실험실) 아래에서 측정했을 때 결과 사이의 일치 정도, (2) **반복성(repeatability)**, 같은 방법과 조건으로 연속 측정했을 때 결과 간의 일치 정도

² 만일 \bar{x} 와 s 가 같은 자료 세트에 근거하고 있다면 자유도는 $n-1$ 이다. 그러나 만일 s 가 부차적인 증거에 근거한다면, 예컨대 관리도에서 제대로 측정된 자료라고 판정을 받았다면 이 때 s 추정에 쓰인 자유도를 t 값 결정에 써도 무방하다. 이런 경우에는 심지어 단 한 점 자료만으로도 신뢰 구간을 구할 수 있다.

1 단계: 검정시 α 의 원하는 확률 수준(유의 수준)을 선택한다.

2 단계: 더 나은 표준 편차 추정을 위해 두 추정치를 한데 합쳐서 계산한다.

$$s_p = \left(\frac{\nu_A s_A^2 + \nu_B s_B^2}{\nu_A + \nu_B} \right)^{1/2} \quad (5)$$

식에서 ν_A 와 ν_B 는 s_A 와 s_B 각각의 자유도이다. s_p 는 $\nu_A + \nu_B$ 의 자유도를 갖는다.

3 단계: 차이의 불확실도 U 를 계산한다.

$$U = t s_p \left(\frac{1}{n_A} + \frac{1}{n_B} \right)^{1/2} \quad (6)$$

식에서 t 는 해당되는 Student t 값이다.

4 단계: 차이의 절대값 $\Delta = |\bar{x}_A - \bar{x}_B|$ 을 U 와 비교한다. 만일 $\Delta \leq U$ 이면 평균에 차이가 난다고 볼 이유가 없다.

경우 2. 표준 편차가 유의하게 다른 경우 (3.5절 참조)

1 단계: 검정시 α 의 원하는 확률 수준(유의 수준)을 선택한다.

2 단계: 각 평균의 표준 편차로 각각의 분산을 계산한다.

$$V_A = s_A^2 / n_A, \quad V_B = s_B^2 / n_B \quad (7)$$

3 단계: 각 통계값의 유효한 자유도를 계산한다³.

$$f^* = \frac{(V_A + V_B)^2}{\frac{V_A^2}{n_A + 1} + \frac{V_B^2}{n_B + 1}} - 2 \quad (8)$$

4 단계: 차이의 불확실도 U 를 계산한다.

$$U = t^* \sqrt{V_A + V_B} \quad (9)$$

t^* 는 자유도 f^* 와 선택한 유의 수준 α 에 맞는 t 의 유효값이다(별첨의 표 4-9 참조).

5 단계: 차이의 절대값 $\Delta = |\bar{x}_A - \bar{x}_B|$ 을 U 와 비교한다. 만일 $\Delta \leq U$ 이면 평균에 차이가 난다고 볼 이유가 없다.

³ 이 계산은 문헌을 보면 여러 가지 식이 등장하는데 몇몇 저자는 분산을 한데 모으는 것(pooling)이 부적절하다고 반박한다. 여기서는 Taylor (1987)의 식을 따랐다.

3.5 표준 편차 추정값의 비교 (F 테스트)

이 검사는 두 표준 편차의 추정값이 유의하게 다른지를 판정하는데 쓰일 수 있다. 두 분산의 비를 구하고 표에 나와있는 값과 비교하는 절차로 되어 있다. 계산 값이 표의 값보다 크지 않다면 표준 편차가 서로 다르다고 볼 이유가 없다.

F 비는 다음과 같이 계산된다.

$$F = \frac{s_L^2}{s_S^2} \quad (10)$$

식에서 s_L 이 두 비교값 가운데 큰 값이고 s_S 가 작은 값이다. F 의 임계값은 유의 수준과 s_L 과 s_S 의 자유도에 따른다 (별첨의 표 4-10 참조).

3.6 최소제곱 추정값 계산

아래 선형 모형이,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (11)$$

x 는 본질적으로 오차가 없고 (x 와 y 에 모두 오차가 있는 경우는 York (1966) 참조) 오차 ε_i 는 상수인 분산을 가지며 정규 분포를 보이면, 최소제곱법으로 추정된 계수 β_0 와 β_1 은 다음 식으로 표현된다.

$$\beta_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} \quad (12)$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \quad (13)$$

이러한 경우 실험 오류의 분산 추정값은 다음과 같이 주어진다.

$$s^2 = \frac{\sum_i (y_i - \beta_0 - \beta_1 x_i)^2}{n-2} \quad (14)$$

그리고 계수에 대한 표준 오차에 대한 추정값은 다음과 같이 주어진다.

$$\text{S.E.}(\beta_0) = s \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_i (x_i - \bar{x})^2} \right)^{1/2} \quad (15)$$

$$\text{S.E.}(\beta_1) = \frac{s}{\left(\sum_i (x_i - \bar{x})^2 \right)^{1/2}} \quad (16)$$

4. 계산 예시

4.1 측정 시리즈의 평균과 표준 편차 추정

아래 9개의 측정값에 대한:

1977.67, 1977.98, 1977.29, 1978.60, 1979.48,
1979.14, 1979.33, 1979.95, 1979.99,

평균은 1978.83이고 표준 편차는 0.99이다.

4.2 반복 측정한 다른 자료 세트에 대한 표준 편차 추정

아래 열 쌍의 측정값에 대해:

1976.8, 1979.3, 1978.9, 1979.6, 1979.6, 1979.8
1978.3, 1978.6, 1981.2, 1979.8, 1977.6, 1977.8
1976.2, 1976.8, 1978.6, 1977.0, 1976.6, 1978.9
1978.3, 1978.9,

아래 식에 따른 표준 편차는

$$s_R = \left(\frac{\sum_{i=1}^k d_i^2}{2k} \right)^{1/2}$$

0.93이다.

4.3 평균 신뢰 구간

4.1 절에서 계산된 평균의 95% 신뢰 구간은

$$1978.83 \pm \frac{(2.306)(0.99)}{\sqrt{9}} = 1978.83 \pm 0.76$$

4.4 두 평균 값의 비교

경우 1. 표준 편차가 다를 이유가 없는 경우.

$$\begin{aligned} \bar{x}_A &= 1978.78, & s_A &= 0.93, & n_A &= 9 \\ \bar{x}_B &= 1981.74, & s_B &= 0.87, & n_B &= 18 \end{aligned}$$

1 단계: 95% 신뢰 수준으로 검정한다.

2 단계: 표준 편차를 한데 모은다.

$$s_p = \left(\frac{8(0.93)^2 + 17(0.87)^2}{8+17} \right)^{1/2}$$

$$= 0.89$$

3단계: U 를 계산한다.

$$U = 2.060(0.89) \left(\frac{1}{9} + \frac{1}{18} \right)^{1/2}$$

$$= 0.75$$

4단계: $\Delta (= 1981.74 - 1978.78 = 2.96)$ 가 U 보다 크므로 95% 신뢰 수준에서 평균은 서로 다르다고 판정한다.

경우 2. 표준 편차가 유의하게 다른 경우

$$\bar{x}_A = 1978.78, \quad s_A = 0.93, \quad n_A = 9$$

$$\bar{x}_B = 1981.74, \quad s_B = 2.75, \quad n_B = 16$$

1 단계: 95% 신뢰 수준으로 검정한다.

2 단계: 각 평균의 분산을 계산한다.

$$V_A = (0.93)^2 / 9 = 0.0961$$

$$V_B = (2.75)^2 / 16 = 0.4727$$

3 단계: 통계값의 유효한 자유도를 계산한다.

$$f^* = \left[\frac{(0.0961 + 0.4727)^2}{(0.0961)^2 / (9 + 1) + (0.4727)^2 / (16 + 1)} \right] - 2 \approx 21$$

4 단계: U 를 계산한다.

$$U = 2.08(0.0961 + 0.4727)^{1/2} = 1.57$$

5 단계: $\Delta (= 1981.74 - 1978.78 = 2.96)$ 가 U 보다 크므로 95% 신뢰 수준에서 평균은 서로 다르다고 판정한다.

4.5 표준 편차 추정값의 비교

$$\bar{x}_A = 1978.78, \quad s_A = 0.93, \quad n_A = 9$$

$$\bar{x}_B = 1975.35, \quad s_B = 1.71, \quad n_B = 12$$

F 를 계산한다:

$$F = \frac{(1.71)^2}{(0.93)^2} = 3.38$$

표에 제시된 F 값은 분자의 자유도 8 과 분모의 자유도 11 에서 3.7 이다. 그런데 계산 결과가 이보다 작으므로 표준 편차가 통계적으로 유의하게 다르다고 볼 이유가 없다.

4.6 최소제곱법 추정값 계산

주어진 x 와 y 여섯 쌍의 자료에 대해

0.0	1892
498.8	66537
1001.9	130818
1500.8	195216
2002.5	260068
2497.1	323456

선형회귀 결과 :

$$\beta_0 = 2017.77$$

$$\beta_1 = 128.765$$

이와 결부된 오차 추정치는 다음과 같다.

$$s = 221.77$$

$$\text{S.E.}(\beta_0) = 160.55$$

$$\text{S.E.}(\beta_1) = 0.106$$

5. 참고 문헌

- Box, G.E.P., Hunter, W.G. and Hunter, J.S. 1978. *Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building*. John Wiley & Sons, 653 pp.
- Meier, P.C. and Zünd, R.E. 1993. *Statistical Methods in Analytical Chemistry*. Wiley Interscience, New York, 321 pp.
- Ryan, T.P. 1989. *Statistical Methods for Quality Improvement*. John Wiley & Sons, Inc., New York, 446 pp.
- Taylor, J.K. 1987. *Quality Assurance of Chemical Measurements*. Lewis Publishers, Inc., Chelsea, 328 pp.
- Taylor, J.K. 1990. *Statistical Techniques for Data Analysis*. Lewis Publishers, Inc., Chelsea, 200 pp.
- York, D. 1966. Least-squares fitting of a straight line. *Can. J. Physics* **44**: 1079–1086.

별첨

표 4-9 신뢰구간 95%와 99%에 대한 Student t 값

양방향 신뢰구간의 확률 수준		
df ⁴	95%	99%
1	12.706	63.657
2	4.303	9.925
3	3.182	5.841
4	2.776	4.604
5	2.571	4.032
6	2.447	3.707
7	2.365	3.499
8	2.306	3.355
9	2.262	3.250
10	2.228	3.169
11	2.201	3.106
12	2.179	3.055
13	2.160	3.012
14	2.145	2.977
15	2.131	2.947
16	2.120	2.921
17	2.110	2.898
18	2.101	2.878
19	2.093	2.861
20	2.086	2.845
25	2.060	2.787
40	2.021	2.704
60	2.000	2.660
∞	1.960	2.576

⁴ 자유도 (degrees of freedom: $n - 1$)

표 4-10 95% 신뢰 구간에서 양쪽꼬리검정 방식으로 등분산을 판정할 때 F 테스트의 임계값.

df_D	df_N									
	1	2	4	6	8	10	15	20	30	40
1	648	800	900	937	957	969	983	993	1001	1006
2	38.5	39.0	39.2	39.3	39.4	39.4	39.4	39.4	39.5	39.5
4	12.2	10.6	9.6	9.2	9.0	8.8	8.7	8.6	8.5	8.4
6	8.8	7.3	6.2	5.8	5.6	5.5	5.3	5.2	5.1	5.0
8	7.6	6.1	5.0	4.6	4.4	4.3	4.1	4.0	3.9	3.8
10	6.9	5.5	4.5	4.1	3.8	3.7	3.5	3.4	3.3	3.3
15	6.2	4.8	3.8	3.4	3.2	3.1	2.9	2.8	2.6	2.6
20	5.9	4.5	3.5	3.1	2.9	2.8	2.6	2.5	2.4	2.3
30	5.6	4.2	3.2	2.9	2.6	2.5	2.3	2.2	2.1	2.0
40	5.4	4.0	3.1	2.7	2.5	2.4	2.2	2.1	1.9	1.9
60	5.3	3.9	3.0	2.6	2.4	2.3	2.1	1.9	1.8	1.7
120	5.2	3.8	2.9	2.5	2.3	2.2	1.9	1.8	1.7	1.6
∞	5.0	3.7	2.8	2.4	2.2	2.1	1.8	1.7	1.6	1.5

df_D — 분모 분산의 자유도

df_N — 분자 분산의 자유도