eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

KIMIKA FAKULTATEA
FACULTAD DE QUÍMICA

2. Gaia. Estatistika. II

Adierazgarritasun irizpideak

Aurkibidea

1. t proba batazbestekoa benetako balio batekin konparatzeko

$$\mu = x \pm \frac{ts}{\sqrt{n}}$$

$$t_{kal} = \left|\left(\bar{X} - \mu\right)\right| \frac{\sqrt{n}}{s}$$

t_{taula} (n−1, %95)

$t_{kal} \le t_{taula} \Rightarrow$ diferentzia ez da adierazgarri da ($x = \mu$).

$t_{kal} > t_{taula} \Rightarrow$ diferentzia adierazgarria da ($x \ne \mu$).

Guk kalkulatutako $\bar{X}$ balioa tarte barruan kokatzen bada esan daiteke gure metodoak ez duela errore sistematikorik, beraien diferentzia errore aleatorioak bakarrik eragiten dutelako.

Erabiltzen den konfiantza maila %95koa da.

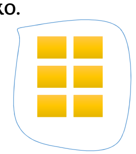
Analisia egin eta neurtu

$\bar{X}$ eta S **konparatu** μ= 105 mg

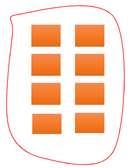
2. t proba batazbesteko esperimentalak konparatzeko

Bi batazbesteko esperimental konparatu nahi direnean lehendabizi batazbesteko horiei dagozkien desbideratze estandarrak konparatu behar dira. Horren arabera, bi prozedura jarrai daitezke batazbestekoak konparatzeko.

Bi datu multzo ditugu: X_1, s_1 eta n_1
X_2, s_2 eta n_2.



X_1, s_1 eta n_1



X_2, s_2 eta n_2

a. Desbideratze estandarrak ezberdinak ez direnean. Desbideratze estandarrak antzekoak direnean

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$t_{kal} = \frac{\left|\bar{X}_1 - \bar{X}_2\right|}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

t_{taula} (n_1+ n_2-2, %95)

$t_{kal} \le t_{taula} \Rightarrow X_1$ eta X_2 arteko diferentzia ez da adierazgarria.

$t_{kal} > t_{taula} \Rightarrow X_1$ eta X_2 arteko diferentzia adierazgarria da.

2. t proba batazbesteko esperimentalak konparatzeko

b. Desbideratze estandarrak ezberdinak direnean.

$$t_{kal} = \frac{\left|\bar{X}_1 - \bar{X}_2\right|}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Askatasun gradoak
Taulan begiratzeko

$$\nu = \left\{ \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 + 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 + 1}} \right\} - 2$$

Konfiantza maila %95

$t_{kal} \le t_{taula} \Rightarrow X_1$ eta X_2 arteko diferentzia ez da adierazgarria.

$t_{kal} > t_{taula} \Rightarrow X_1$ eta X_2 arteko diferentzia adierazgarria da.

3. F proba bariantzak konparatzeko

$$F_{kal} = \frac{S_1^2}{S_2^2}$$ non $s_1 > s_2$ BETI >1,0

$F_{kal} \le F_{taula}$ ⇨ diferentzia ez da adierazgarria da ($S_1 = S_2$).

$F_{kal} > F_{taula}$ ⇨ diferentzia adierazgarria da ($S_1 \ne S_2$).

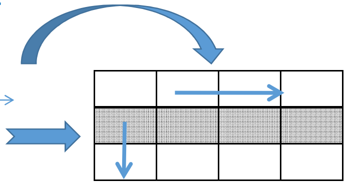
F_{taula} (n_1-1,n_2-1,%95)

Hurrengo taulan begiratzeko:

$$F_{kal} = \frac{S_1^2}{S_2^2}$$

S_1 ⇨⇨ ν_1

S_2 ⇨⇨ ν_2



Adibidez: $\nu_1 = n_1 - 1$ $\quad n_1 = 5$ → F= 4.12

$\nu_2 = n_2 - 1$ $\quad n_2 = 8$

2. F proba bariantzak konparatzeko

Table 4-5 Critical values of $F = s_1^2/s_2^2$ at 95% confidence level

Degrees of freedom for s_2	Degrees of freedom for s_1: 2	3	4	5	6	7	8	9	10	12	15	20	30	∞
2	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5
3	9.55	9.28	9.12	9.01	8.94	8.89	8.84	8.81	8.79	8.74	8.70	8.66	8.62	8.53
4	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.63
5	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.36
6	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.67
7	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.58	3.51	3.44	3.38	3.23
8	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	2.93
9	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.71
10	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.84	2.77	2.70	2.54
11	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.57	2.40
12	3.88	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.30
13	3.81	3.41	3.18	3.02	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.38	2.21
14	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.31	2.13
15	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.07
16	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.19	2.01
17	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.15	1.96
18	3.56	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.11	1.92
19	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.07	1.88
20	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.84
30	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.62
∞	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.00

4. Binakako laginen konparaketa

Hipotesia: bi metodoen artean ezberdintasunik ez balego, emaitzen arteko ezberdintasunak m=0 batazbestekoa duen populazioari dagokio.

Hipotesi hau frogatzeko binakako datuen diferentzien arteko batazbesteko esperimentala ($\bar{d}$) konparatzen da 0 balioarekin *t* proba bat eginez.

$$\boxed{t_{kal} = \frac{\bar{d}\sqrt{n}}{s_d}}$$

$\bar{d}$: binakako balioen arteko diferentzien batazbestekoa

S_d : binakako balioen arteko diferentzien desbideratze estandarra

Konparatu t_{taula} (n-1, %95)

4. Binakako laginen konparaketa

Bi metodo desberdin erabiltzen parazetamola determinatu da pilula ezberdinetan

Parazetamola (%)

Laginak	UV espektroscopia	NIR espektroskopia	d
1	84.63	83.15	1.48
2	84.38	83.72	0.66
3	84.08	83.84	0.24
4	84.41	84.20	0.21
5	83.82	83.92	-0.10

Adierazgarria da bi metodoen arteko ezberdintasuna?

$\bar{d}$ = 0.498

S_d = 0.612

$$t_{kal} = \frac{\bar{d}\sqrt{n}}{s_d} = \frac{0.498\sqrt{5}}{0.612} = \frac{1.114}{0.612} = 1.820$$

t_{taula}= 2.776 $\quad t_{kal} < t_{taula}$ diferentzia ez da adierazgarria

Dixon-en Q

$$Q_{kal} = \frac{|Balore\ susmagarria - Balore\ hurbilena|}{Balore\ handiena - Balore\ txikiena}$$

Q_{taula} (n, %95)

Table 4-6 Values of Q for rejection of data

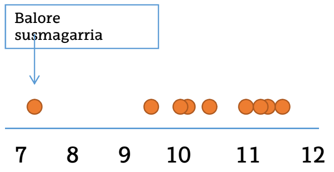
Q (90% confidence)[a]	Number of observations
0.76	4
0.64	5
0.56	6
0.51	7
0.47	8
0.44	9
0.41	10

a. Q = gap/range. If $Q_{calculated} > Q_{table}$, the value in question can be rejected with 90% confidence.

SOURCE: R. B. Dean and W. J. Dixon, *Anal. Chem.* **1951**, *23*, 636; see also D. R. Rorabacher, *Anal. Chem.* **1991**, *63*, 139.

$Q_{kal} \leq Q_{taula}$ ⇨ Balore susmagarria mantendu.

$Q_{kal} > Q_{taula}$ ⇨ Balore susmagarria eliminatu.



5. Emaitza susmagarriak baztertzeko probak

Dixon-en Q

Adibidea: Hemen dituzu 10 datu esperimentalak, saio batean lortuak:

10.45; 10.47; 10.47; 10.48; 10.49; 10.50; 10.50; 10.52; 10.53; 10.58

10.58 datu susmagarria dela esan genezake?

$$Q_{kal} = \frac{|Balore\ susmagarria - Balore\ hurbilena|}{Balore\ handiena - Balore\ txikiena} = \frac{10.58 - 10.53}{10.58 - 10.45} = \frac{0.05}{0.13} = 0.385$$

%95eko konfiantza maila kontutan harturik,

$Q_{taula} = 0.41$ beraz $Q_{kal} < Q_{taula}$

balioa ez da susmagarria beraz mantendu

Ariketa batean balio susmagarria baztertu behar bada, baztertu eta birkalkulatu X eta S

5. Emaitza susmagarriak baztertzeko probak

Grubbs-en G

$$G_{kalk} = \frac{|Balore\ susmagarria - \bar{X}|}{s}$$

G_{taula} (n, %95)

$G_{kal} \leq G_{taula}$ ⇨ Balore susmagarria mantendu.

$G_{kal} > G_{taula}$ ⇨ Balore susmagarria eliminatu.

TABLE 4-5 Critical values of G for rejection of outlier

Number of observations	G (95% confidence)
4	1.463
5	1.672
6	1.822
7	1.938
8	2.032
9	2.110
10	2.176
11	2.234
12	2.285
15	2.409
20	2.557

$G_{calculated}$ = \|questionable value − mean\|/s. If $G_{calculated} > G_{table}$, the value in question can be rejected with 95% confidence. Values in this table are for a one-tailed test, as recommended by ASTM.

SOURCE: ASTM E 178-02 Standard Practice for Dealing with Outlying Observations, http://webstore.ansi.org; F. E. Grubbs and G. Beck, Technometrics 1972, 14, 847.

X eta S kalkulatzeko balio susmagarria erabili

Ariketa batean balio susmagarria baztertu behar bada, baztertu eta birkalkulatu X eta S

5. Emaitza susmagarriak baztertzeko probak

Grubbs-en G

Adibidea: Hemen dituzu 12 datu esperimentalak, saio batean lortuak:

10.2; 10.8; 10.6; 9.9; 9.4; 6.8; 10.0; 9.2; 11.3; 9.5; 10.6; 11.2

6.8 datu susmagarria dela esan genezake?

Lehendabizi X eta s kalkulatu, datu guztiak kontutan hartzen.

X=9,96 s=1.21 n=12

$$G_{kal} = \frac{|Balore\ susmagarria - \bar{X}|}{s} = \frac{|6.8 - 9.96|}{1.21} = \frac{3.16}{1.21} = 2.612$$

%95eko konfiantza maila kontutan harturik,

$G_{taula} = 2.285$ beraz $G_{kal} > G_{taula}$

balioa susmagarria da beraz baztertu

Ondorioz, birkalkulatu X eta S. ⇨ X=10,24 s=0,72 n=11