

Za rešavanje zadataka potrebno je koristiti formule i tabele date na trećem predavanju.

1. U toku vežbe na predmetu Opšti kurs fizičke hemije, student je imao zadatak da piknometrom izmeri gustinu rastvora. Merenje gustine rastvora je ponovljeno deset puta i dobijene su sledeće vrednosti (u gcm^{-3}):

1,06	1,12	1,15	1,06	1,03	1,09	1,22	1,08	1,13	1,02
------	------	------	------	------	------	------	------	------	------

Izračunati srednju vrednost, standardnu devijaciju, relativnu standardnu devijaciju, modu, opseg merenja, varijansu i medijanu.

2. Na osnovu rezultata iz prethodnog merenja gustine rastvora piknometrom pokazati Grubbs-ovim testom da li je moguće isključiti vrednost $1,22 \text{ gcm}^{-3}$ sa nivoom pouzdanosti od 95%.
3. Na osnovu rezultata iz prethodnog merenja gustine rastvora piknometrom pokazati Dixon-ovim testom da li je moguće isključiti vrednost $1,22 \text{ gcm}^{-3}$ sa nivoom pouzdanosti od 95%. Da li je na osnovu istog testa moguće isključiti vrednost $1,02 \text{ gcm}^{-3}$.
4. Za merenje gustine rastvora piknometrom izračunati opseg intervala pouzdanosti od 50, 90 i 95%.
5. Većim brojem merenja tokom godina izračunato je da je standardna devijacija svih merenja gustine datog rastvora piknometrom $0,05 \text{ gcm}^{-3}$. Pokazati F-testom da li između merenja i populacije postoji statistički značajna razlika
6. Ukoliko je na pakovanju ispitivanog rastvora pisalo da je gustina $1,16 \text{ gcm}^{-3}$ pokazati da li postoji statistički značajna razlika između ove vrednosti i srednje vrednosti merenja.

7. Drugi student je ponovio merenje i dobio sledeći skup vrednosti (u gcm^{-3}):

1,18	1,16	1,20	1,25	1,15	1,13	1,28	1,16	1,13	1,55
------	------	------	------	------	------	------	------	------	------

Pokazati da li postoji statistički značajna razlika između dva merenja ukoliko se uporede srednje vrednosti merenja.

8. Pokazati da li postoji statistički značajna razlika između merenja dvoje studenata ukoliko se uporede varijanse.

2. ДОМАЋИ ЗАДАТАК

РЕШЕЊА

1. задатак

Вредности густине дате у задатку											
i	1	2	3	4	5	6	7	8	9	10	suma
$\rho_i \text{ (g cm}^{-3}\text{)}$	1,06	1,12	1,15	1,06	1,03	1,09	1,22	1,08	1,13	1,02	10,96

$$\rho_{sr} = \frac{\sum_{i=1}^{10} \rho_i}{10} = \frac{10,96}{10} = 1,096 \text{ g} \cdot \text{cm}^{-3} \text{ - средња вредност густине}$$

Подаци за густину разврстани по растућим вредностима за одређивање медијане										
i	1	2	3	4	5	6	7	8	9	10
$\rho_i \text{ (g cm}^{-3}\text{)}$	1,02	1,03	1,06	1,06	1,08	1,09	1,12	1,13	1,15	1,22

Пошто је број мерења паран, медијана се одређује тако што се две средишње вредности саберу и поделе са два. У овом случају, у питању су 5. и 6. вредност.

$$\text{медијана} = (1,08 + 1,09)/2 \text{ g cm}^{-3} = 1,085 \text{ g cm}^{-3}$$

Мода је вредност која се појављује највећи број пута = 1,06 g cm⁻³

Опсег је распон вредности од најмање до највеће вредности.

$$\text{опсег} = (1,22 - 1,02) \text{ g cm}^{-3} = 0,20 \text{ g cm}^{-3}$$

Поступак одређивања стандардне девијације											
$\rho_i \text{ (g cm}^{-3}\text{)}$	1,06	1,12	1,15	1,06	1,03	1,09	1,22	1,08	1,13	1,02	10,96
$ \rho_{sr} - \rho_i \text{ (g cm}^{-3}\text{)}$	0,036	0,024	0,054	0,036	0,066	0,006	0,124	0,016	0,034	0,076	0,472
$(\rho_{sr} - \rho_i)^2 \text{ (g}^2 \text{ cm}^{-6}\text{)}$	0,001296	0,000576	0,002916	0,001296	0,004356	3,6E-05	0,015376	0,000256	0,001156	0,005776	0,03304

$$s = \sqrt{\frac{\sum_{i=1}^{10} (\rho_{sr} - \rho_i)^2}{10-1}} = \sqrt{\frac{0,03304}{9} \text{ g}^2 \cdot \text{cm}^{-6}} = 0,060589695 \text{ g} \cdot \text{cm}^{-3} \text{ - стандардна девијација мерења}$$

$$\text{Релативна стандардна девијација } s_R = \frac{s}{\rho_{sr}} \cdot 100\% = \frac{0,0606 \text{ g} \cdot \text{cm}^{-3}}{1,096 \text{ g} \cdot \text{cm}^{-3}} \cdot 100\% = 5,53\%$$

$$\text{Варијанса } s^2 = 0,0037$$

2. задатак

Сумњива/упитна вредност $1,22 \text{ g cm}^{-3}$

Испитати помоћу Грубсовог теста

ниво поузданости 95%

број мерења 10

број степени слободе 9

На основу нивоа поузданости и броја мерења, параметар G из таблице је $G_{\text{tab}} = 2,176$.

$$G_{\text{izr}} = \frac{|\text{упитна вредност} - \rho_{\text{sr}}|}{s} = \frac{|1,22 - 1,096|}{0,0606} = 2,046$$

$G_{\text{izr}} < G_{\text{tab}}$ израчуната вредност параметра G је мања од табличне вредности. Упитна вредност густине од $1,22 \text{ g cm}^{-3}$ није аутлејер и може се задржати.

Вредности параметра G за одбацивање аутлејера	
Број мерења	Интервал поузданости (%)
	95
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

3. задатак

Сумњиве/упитне вредности $1,22$ и $1,02 \text{ g cm}^{-3}$

Испитати помоћу Грубсовог теста

ниво поузданости 95%

број мерења 10

број степени слободе 9

На основу нивоа поузданости и броја мерења, параметар Q из таблице је $Q_{\text{tab}} = 0,466$.

Q израчунато за већу упитну вредност:

$$Q_{\text{izr}} (\text{већа}) = \frac{\text{већа упитна вредност} - \rho_9}{\rho_{10} - \rho_1} = \frac{1,22 - 1,15}{1,22 - 1,02} = 0,35$$

Q израчунато за мању упитну вредност:

$$Q_{\text{izr}} (\text{мању}) = \frac{\rho_2 - \text{мања упитна вредност}}{\rho_{10} - \rho_1} = \frac{1,03 - 1,02}{1,22 - 1,02} = 0,05$$

$Q_{\text{izr}} < Q_{\text{tab}}$ израчуната вредност параметра Q је мања од табличне вредности за обе упитне вредности. Упитне вредности густине нису аутлејери и не могу се искључити.

Вредности параметра Q за одбацивање аутлејера	
Број мерења	Интервал поузданости (%)
	95
3	0,970
4	0,829
5	0,710
6	0,625
7	0,568
8	0,526
9	0,493
10	0,466

4. задатак

$$\text{Опсег интервала поузданости } \rho_{\text{sr}} \pm \frac{t_{\text{tab}} \cdot s}{\sqrt{n}}$$

ниво поузданости 50%, 90% и 95%

број мерења 10

број степени слободе 9

За читање табличне вредности параметра t гледа се број степени слободе и ниво поузданости. За број степени слободе 9, табличне вредности за параметар t су:

$$t_{50} = 0,703 \quad t_{90} = 1,833 \quad t_{95} = 2,262$$

опсег за ниво поузданости 50%

$$\left(1,096 \pm \frac{0,703 \cdot 0,0606}{\sqrt{10}} \right) \text{g} \cdot \text{cm}^{-3} = (1,096 \pm 0,013) \text{g} \cdot \text{cm}^{-3} = (1,10 \pm 0,02) \text{g} \cdot \text{cm}^{-3}$$

опсег за ниво поузданости 90%

$$\left(1,096 \pm \frac{1,833 \cdot 0,0606}{\sqrt{10}} \right) \text{g} \cdot \text{cm}^{-3} = (1,096 \pm 0,035) \text{g} \cdot \text{cm}^{-3} = (1,10 \pm 0,04) \text{g} \cdot \text{cm}^{-3}$$

опсег за ниво поузданости 95%

$$\left(1,096 \pm \frac{2,262 \cdot 0,0606}{\sqrt{10}} \right) \text{g} \cdot \text{cm}^{-3} = (1,096 \pm 0,043) \text{g} \cdot \text{cm}^{-3} = (1,10 \pm 0,05) \text{g} \cdot \text{cm}^{-3}$$

Вредности параметра t за одбацивање аутлејера			
Број мерења	Интервал поузданости (%)		
	50	90	95
1	1,000	6,314	12,706
2	0,816	2,920	4,302
3	0,765	2,353	3,182
4	0,741	2,132	2,776
5	0,727	2,015	2,571
6	0,718	1,943	2,447
7	0,711	1,895	2,365
8	0,706	1,860	2,306
9	$t_{50} = 0,703$	$t_{90} = 1,833$	$t_{95} = 2,262$
10	0,700	1,812	2,228
15	0,691	1,753	2,131
20	0,687	1,725	2,086
25	0,684	1,708	2,060
30	0,683	1,697	2,042
40	0,681	1,684	2,021
60	0,679	1,671	2,000
120	0,677	1,658	1,980
∞	0,674	1,645	1,960

5. задатак

Ниво поузданости 95%

Стандардна девијација популације је $s_{\text{pop}} = 0,05 \text{ g cm}^{-3}$

Варијанса за популацију је $(s_{\text{pop}})^2 = 0,0025 \text{ g}^2 \text{ cm}^{-6}$

Број степени слободе популације је ∞ .

Стандардна девијација узорка је $s_{\text{uz}} = 0,05 \text{ g cm}^{-3}$

Варијанса за популацију је $(s_{\text{uz}})^2 = 0,00367 \text{ g}^2 \text{ cm}^{-6}$

Број степени слободе узорка је 9.

$$F_{\text{izr}} = \frac{\text{већа варијанса}}{\text{мања варијанса}} = \frac{(s_{\text{uz}})^2}{(s_{\text{pop}})^2} = \frac{0,00367}{0,0025} = 1,468 \text{ израчуната вредност параметра } F.$$

Табличне F вредности за ниво поузданости од 95% (вредности v_1 и v_2 представљају број степени слободе за серију мерења са већом и са мањом варијансом, респективно.

v_1/v_2	1	2	3	4	5	6	7	8	9	10	15	20	∞
1	647,8	799,5	846,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	984,9	993,1	1018
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,43	39,45	39,498
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,25	14,17	13,902
4	12,22	10,65	9,979	9,605	9,364	9,197	9,074	8,980	8,905	8,844	8,657	8,560	8,257
5	10,01	8,434	7,764	7,388	7,146	6,978	6,853	6,757	6,681	6,619	6,428	6,329	6,015
6	8,813	7,260	6,599	6,277	5,988	5,820	5,695	5,600	5,523	5,461	5,269	5,168	4,849
7	8,073	6,542	5,890	5,523	5,285	5,119	4,995	4,899	4,823	4,761	4,568	4,467	4,142
8	7,571	6,059	5,416	5,053	4,817	4,652	4,529	4,433	4,357	4,295	4,101	3,999	3,670
9	7,209	5,715	5,078	4,718	4,484	4,320	4,197	4,102	4,026	3,964	3,769	3,667	3,333
10	6,937	5,456	4,826	4,468	4,236	4,072	3,950	3,855	3,779	3,717	3,522	3,419	3,080
11	6,724	5,256	4,630	4,275	4,044	3,881	3,759	3,664	3,588	3,526	3,330	3,226	2,883
12	6,544	5,096	4,474	4,121	3,891	3,728	3,607	3,512	3,436	3,374	3,177	3,073	2,725
13	6,414	4,965	4,347	3,996	3,767	3,604	3,483	3,388	3,312	3,250	3,053	2,948	2,596
14	6,298	4,857	4,242	3,892	3,663	3,501	3,380	3,285	3,209	3,147	2,949	2,844	2,487
15	6,200	4,765	4,153	3,804	3,576	3,415	3,293	3,199	3,123	3,060	2,862	2,756	2,395
16	6,115	4,687	4,077	3,729	3,502	3,341	3,219	3,125	3,049	2,986	2,788	2,681	2,316
17	6,042	4,619	4,011	3,665	3,438	3,277	3,156	3,061	2,985	2,922	2,723	2,616	2,247
18	5,978	4,560	3,965	3,608	3,382	3,221	3,100	3,005	2,929	2,866	2,667	2,559	2,187
19	5,922	4,508	3,903	3,559	3,333	3,172	3,051	2,956	2,880	2,817	2,617	2,509	2,133
20	5,871	4,461	3,859	3,515	3,289	3,128	3,007	2,913	2,837	2,774	2,573	2,464	2,085
∞	5,024	3,689	3,116	2,786	2,567	2,408	2,288	2,192	2,114	2,048	1,833	1,708	1,000

Приликом читавања табличне вредности параметра F, F_{tab} гледа се број степени слободе за обе серије мерења и за популацију и за узорак.

У првој колони су вредности степени слободе за серију мерења са већом варијансом, а у првом реду су вредности степени слободе са мањом варијансом. У пресеку степени слободе обе серије мерења налази се таблична вредност F_{tab} .

У нашем случају узорак има већу варијансу од популације. У првом реду се гледа број степени слободе 9, а у првој колони се гледа број степени слободе ∞ . У његовом пресеку је тражена таблична вредност, $F_{tab} = 2,114$.

$$F_{izr} < F_{tab}$$

Пошто је израчуната вредност параметра F мања од табличне, може се сматрати да не постоји статистички значајна разлика између две серије мерења.

6. задатак

оčekивана вредност за густина $\mu = 1,16 \text{ g cm}^{-3}$

средња вредност наших мерења густине је $\rho_{sr} = 1,096 \text{ g cm}^{-3}$

стандардна девијација је $0,0606 \text{ g cm}^{-3}$

број степени слободе је 9

$$t_{izr} = \frac{|\mu - \rho_{sr}| \cdot \sqrt{n}}{s} = \frac{|1,16 - 1,096| \text{ g} \cdot \text{cm}^{-3} \cdot \sqrt{10}}{0,0606 \text{ g} \cdot \text{cm}^{-3}} = 3,34$$

У задатку 4 смо видели да таблична вредност за параметар t за ниво поузданости 95% и број степени слободе 9 износи: $t_{tab} = 2,262$

$$t_{izr} > t_{tab}$$

Пошто је израчуната вредност параметра t већа од табличне, постоји статистички значајна разлика између очекиване вредности и средње вредности серије мерења.

7. задатак

Вредности густине дате за другу серију мерења											
i	1	2	3	4	5	6	7	8	9	10	suma
$\rho_{i,2} \text{ (g cm}^{-3}\text{)}$	1,18	1,16	1,20	1,25	1,15	1,13	1,28	1,16	1,13	1,55	12,19

$$\rho_{sr,2} = \frac{\sum_{i=1}^{10} \rho_{i,2}}{10} = \frac{12,19}{10} = 1,219 \text{ g} \cdot \text{cm}^{-3} \text{ средња вредност за другу серију мерења.}$$

Поступак одређивања стандардне девијације за другу серију мерења											
$\rho_{i,2} \text{ (g cm}^{-3}\text{)}$	1,18	1,16	1,2	1,25	1,15	1,13	1,28	1,16	1,13	1,55	12,19
$ \rho_{sr,2} - \rho_{i,2} \text{ (g cm}^{-3}\text{)}$	0,039	0,059	0,019	0,031	0,069	0,089	0,061	0,059	0,089	0,331	0,846
$(\rho_{sr,2} - \rho_{i,2})^2 \text{ (g}^2 \text{ cm}^{-6}\text{)}$	0,001521	0,003481	0,000361	0,000961	0,004761	0,007921	0,003721	0,003481	0,007921	0,109561	0,14369

$$s_2 = \sqrt{\frac{\sum_{i=1}^{10} (\rho_{sr,2} - \rho_{i,2})^2}{10-1}} = \sqrt{\frac{0,14369}{9} \text{ g}^2 \cdot \text{cm}^{-6}} = 0,1264 \text{ g} \cdot \text{cm}^{-3} \text{ стандардна девијација за другу серију мерења.}$$

Прва серија мерења је дата у задатку 1.

Средња вредности за прву серију мерења $\rho_{sr,1} = 1,096 \text{ g cm}^{-3}$

стандардна девијација за прву серију мерења $s_1 = 0,0606 \text{ g cm}^{-3}$

Средња вредности за прву серију мерења $\rho_{sr,2} = 1,219 \text{ g cm}^{-3}$

стандардна девијација за прву серију мерења $s_2 = 0,1264 \text{ g cm}^{-3}$

број мерења за сваку серију је 10

број степени слободе за сваку серију је 9

ниво поузданости 95%

За две серије мерења заједнички параметар t_{izr} се добија према формули:

$$t = \frac{|\rho_{sr,1} - \rho_{sr,2}|}{s_{zajednicko}} \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} = \frac{|1,096 \text{ g cm}^{-3} - 1,219 \text{ g cm}^{-3}|}{0,09912 \text{ g} \cdot \text{cm}^{-3}} \cdot \sqrt{\frac{10 \cdot 10}{10 + 10}} = 2,775$$

Заједничка вредност стандардне девијације за обе серије мерења добија се преко обрасца:

$$s_{zajednicko} = \sqrt{\frac{\sum_{i=1}^{10} (\rho_{i,1} - \rho_{sr,1})^2 + \sum_{i=1}^{10} (\rho_{i,2} - \rho_{sr,2})^2}{n_1 + n_2 - 2}} = \sqrt{\frac{s_1^2 \cdot (n_1 - 1) + s_2^2 \cdot (n_2 - 1)}{n_1 + n_2 - 2}}$$

$$s_{zajednicko} = \sqrt{\frac{0,00367 \text{ g}^2 \cdot \text{cm}^{-6} \cdot (10 - 1) + 0,01598 \text{ g}^2 \cdot \text{cm}^{-6} \cdot (10 - 1)}{10 + 10 - 2}} = 0,09912 \text{ g} \cdot \text{cm}^{-3}$$

Таблична вредност параметра t за ниво поузданости 95% и број степени слободе 18 није приказана у табlici, али за број степени слободе је 2,228 и за број степени слободе је 2,131. У оба случаја израчуната вредност параметра t_{izr} је мања од табличне. Закључак је да не постоји статистички значајна разлика између две серије мерења.

8. задатак

стандардна девијација за прву серију мерења $s_1 = 0,0606 \text{ g cm}^{-3}$

стандардна девијација за прву серију мерења $s_2 = 0,1264 \text{ g cm}^{-3}$

број мерења за сваку серију је 10

број степени слободе за сваку серију је 9

ниво поузданости 95%

$$F_{izr} = \frac{\text{већа варијанса}}{\text{мања варијанса}} = \frac{(s_2)^2}{(s_1)^2} = \frac{0,01598}{0,00367} = 4,354$$

Таблична вредност када обе серије имају број степени слободе 9 и за ниво поузданости 95% је 4,026.

Израчуната вредност параметра F је већа од табличне. Закључак је да постоје статистички значајне разлике када се упореде две серије мерења преко њихових варијанси.