תאריך הבחינה:	4.4.2024
המרצה:	דר' יקיר ברצ'נקו
שם הקורס:	מודלים של רגרסיה ליניארית
מספר קורס:	364-1-1061
משך הבחינה:	שתי שעות
חומר עזר:	מחשבון, 5 דפי נוסחאות, טבלאות סטטיסטיות

מיועד לתלמידי הנדסת תעשיה וניהול.

שנה: 2024 סמסטר: א' מועד: א'

הנחיות חשובות:

א. במבחן זה שלוש שאלות

ב. יש לפתור את כל השאלות, סך הנקודות עשוי להגיע ל108, אך הציון הסופי המקסימלי במבחן הינו 100.

ג. יש ללוות את דרך הפתרון בהסברים מילוליים קצרים

ד. ניתן להשתמש בכל הנוסחאות והתוצאות שפותחו או הוצגו בכיתה ואין צורך להוכיחן מחדש אלא אם כן זה נדרש במפורש

יש לכתוב את כל התשובות במחברת ולא על גיליון המבחן.

<u>שאלה 1 (45 נקודות)</u>

כימאים אספו נתונים לגבי תגובה כימית, ולאחר הרצת רגרסיה לינארית הם קיבלו את התוצאות הבאות עבור המודל:
(שימו לב שחלק מהערכים בפלט נמחקו!)

```
Call:
lm(formula = chem.data$y ~ chem.data$x1 + chem.data$x2 + chem.data$x3 +
    chem.data$x4)

Residuals:
    Min      1Q  Median      3Q     Max
-9.9958 -3.3092 -0.2419  3.3924 10.5668

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.89453    4.32508   ???     ???
chem.data$x1 -0.47790    0.34002   ???     ???
chem.data$x2  0.18271    0.01718   ???     ???
chem.data$x3 35.40284   11.09960   ???     ???
chem.data$x4  5.84391    2.90978   ???     ???
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.014 on 57 degrees of freedom
Multiple R-squared:  0.6914,   Adjusted R-squared:  0.6697
F-statistic: 31.92 on 4 and 57 DF,  p-value: ???
```

א. האם מודל הרגרסיה מובהק? הסבירו את תשובתכם.

ב. חשבו את הp-value עבור המקדם של x4. הסבירו את תשובתכם.

ג. חשבו רווח סמך ברמה של 99% עבור המקדם של x2. הסבירו את תשובתכם.

שאלה 2 (45 נקודות)

באיורים 10.12 - 10.19 למטה מוצגות התוצאות של ניתוח שאריות לאחר ביצוע רגרסיה מרובה עם נתוני כבישי אספלט.

א. עבור כל איור ציינו והסבירו איזה סימני-אזהרה הוא מעלה לגבי הנחות המודל (בשתיים-שלוש שורות). אם אין כאלו, כיתוב "אין מה לציין" (5 נקודות עבור כל איור, סה"כ 40 נק'). הערה 1: בחלק מהאיורים התשובה איננה חד-משמעית, ויותר חשוב לנו לראות את הנימוקים ועקביות עם התשובות האחרות שכתבתם.

הערה 2: אם ציינתם באיור מסוים שיש בו בעיות מסוימות, כשלמעשה אין בו **שום** עדות לכך, יתכן שירד ניקוד על כך!

ב. כיצד הייתם מטפלים בבעיות השונות שציינתם? 5 נק'.

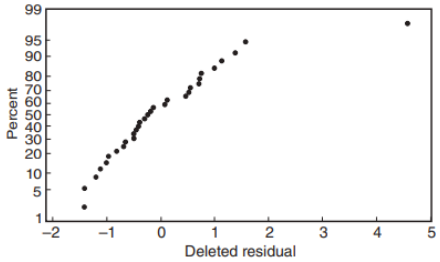


Figure 10.12 Normal probability plot of the residuals for the asphalt data.

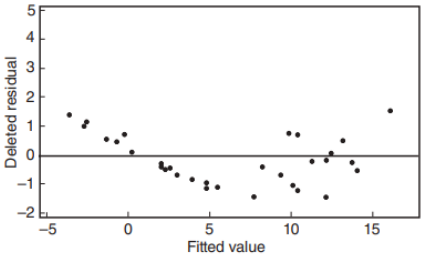


Figure 10.13 Residuals versus the fitted values for the asphalt data.

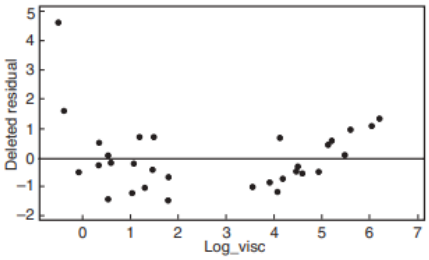


Figure 10.14 Residuals versus the log of the viscosity for the asphalt data.

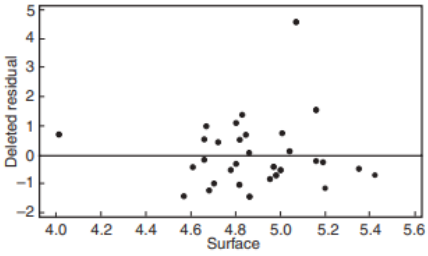


Figure 10.15 Residuals versus surface for the asphalt data.

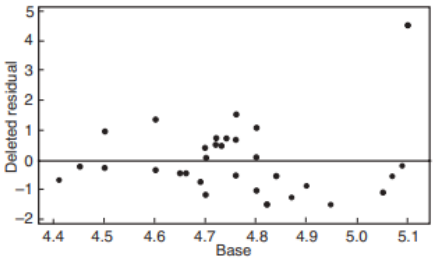


Figure 10.16 Residuals versus base for the asphalt data.

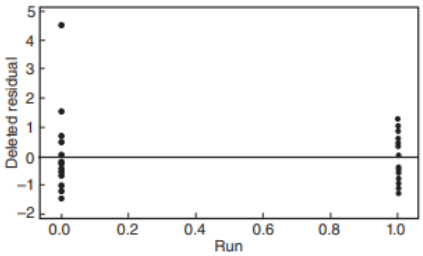


Figure 10.17 Residuals versus run for the asphalt data.

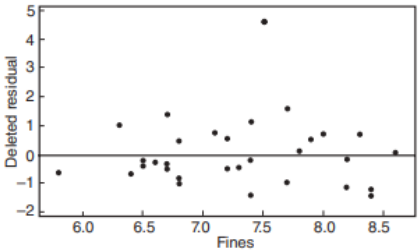


Figure 10.18 Residuals versus fines for the asphalt data.

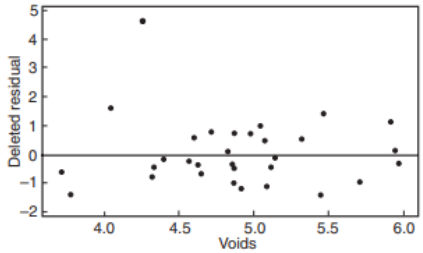


Figure 10.19 Residuals versus voids for the asphalt data.

שאלה 3 (18 נקודות)

יהי $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad i = 1, 2, 3 \dots n$
וכרגיל, ε_i מתפלג נורמלית עם תוחלת 0 ושונות σ_ε^2.

תמי רוצה לבצע רגרסיה כדי למצוא את הקשר בין x ל-y. לצערה של תמי, היא לא יכולה למדוד את ערכי ה-x באופן מדויק, ומקום זאת היא יכולה למדוד רק "גירסה רועשת" שלהם:

$$z_i = x_i + \delta_i$$

כאשר δ_i הינו משתנה מקרי שמתפלג נורמלית עם תוחלת 0 ושונות σ_δ^2.

תמי, שאיננה יכולה לדעת את ערכי x האמיתיים (וגם לא את ערכי δ), טוענת:

"מכיוון δ_i הינו משתנה מקרי חסר-הטיה בעל תוחלת 0, הרי שזה כאילו מוסיפים שגיאה חסרת-הטיה ל-ε . בהתאם, ניתן להשתמש בערכי z במקום ערכי x ולמצוא כרגיל את אומדי הריבועים הפחותים עבור המודל

$$y_i = \beta_0 + \beta_1 z_i + \tilde{\varepsilon}_i \qquad i = 1, 2, 3 \dots n$$

והאומדים שנקבל כך יהיו הערכים הרצויים (עבור מדגם גדול לפחות)"

א. עבור $\widehat{\beta_1}$, אומד הריבועים הפחותים לשיפוע, האם תמי צודקת? בפרט, מה יקרה עבור גודל מדגם גדול (אם תמי תשתמש בערכי z במקום ערכי x)?
הדרכה: סמנו את השונות של x בעזרת σ_x^2 (אליה השונות המדגמית של x מתכנסת), ניתן להעזר גם במשפט סלוצקי (ראו למטה).

ב. תנו דוגמה והסבירו מתי/איך עשויים לדעת את σ_δ^2 , או לקבל אומדן שלו. מה תמליצו לתמי לעשות במקרה כזה?

משפט סלוצקי (חלקי):
יהי $\{a_n\}$ ו- $\{b_n\}$ סדרות של משתנים מקריים.
אם a_n מתכנס בהתפלגות ל-A, ו-b_n מתכנס בהסתברות לקבוע b (אשר שונה מ-0), אזי:

$$\frac{a_n}{b_n} \longrightarrow \frac{A}{b}$$

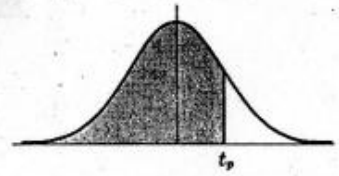
(כאשר ההתכנסות הינה בהתפלגות).

TABLE

PERCENTILE VALUES (t_p) FOR STUDENT'S t DISTRIBUTION

with n degrees of freedom

(shaded area = p)



n	$t_{.995}$	$t_{.99}$	$t_{.975}$	$t_{.95}$	$t_{.90}$	$t_{.80}$	$t_{.75}$	$t_{.70}$	$t_{.60}$	$t_{.55}$
1	63.66	31.82	12.71	6.31	3.08	1.376	1.000	.727	.325	.158
2	9.92	6.96	4.30	2.92	1.89	1.061	.816	.617	.289	.142
3	5.84	4.54	3.18	2.35	1.64	.978	.765	.584	.277	.137
4	4.60	3.75	2.78	2.13	1.53	.941	.741	.569	.271	.134
5	4.03	3.36	2.57	2.02	1.48	.920	.727	.559	.267	.132
6	3.71	3.14	2.45	1.94	1.44	.906	.718	.553	.265	.131
7	3.50	3.00	2.36	1.90	1.42	.896	.711	.549	.263	.130
8	3.36	2.90	2.31	1.86	1.40	.889	.706	.546	.262	.130
9	3.25	2.82	2.26	1.83	1.38	.883	.703	.543	.261	.129
10	3.17	2.76	2.23	1.81	1.37	.879	.700	.542	.260	.129
11	3.11	2.72	2.20	1.80	1.36	.876	.697	.540	.260	.129
12	3.06	2.68	2.18	1.78	1.36	.873	.695	.539	.259	.128
13	3.01	2.65	2.16	1.77	1.35	.870	.694	.538	.259	.128
14	2.98	2.62	2.14	1.76	1.34	.868	.692	.537	.258	.128
15	2.95	2.60	2.13	1.75	1.34	.866	.691	.536	.258	.128
16	2.92	2.58	2.12	1.75	1.34	.865	.690	.535	.258	.128
17	2.90	2.57	2.11	1.74	1.33	.863	.689	.534	.257	.128
18	2.88	2.55	2.10	1.73	1.33	.862	.688	.534	.257	.127
19	2.86	2.54	2.09	1.73	1.33	.861	.688	.533	.257	.127
20	2.84	2.53	2.09	1.72	1.32	.860	.687	.533	.257	.127
21	2.83	2.52	2.08	1.72	1.32	.859	.686	.532	.257	.127
22	2.82	2.51	2.07	1.72	1.32	.858	.686	.532	.256	.127
23	2.81	2.50	2.07	1.71	1.32	.858	.685	.532	.256	.127
24	2.80	2.49	2.06	1.71	1.32	.857	.685	.531	.256	.127
25	2.79	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
26	2.78	2.48	2.06	1.71	1.32	.856	.684	.531	.256	.127
27	2.77	2.47	2.05	1.70	1.31	.855	.684	.531	.256	.127
28	2.76	2.47	2.05	1.70	1.31	.855	.683	.530	.256	.127
29	2.76	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
30	2.75	2.46	2.04	1.70	1.31	.854	.683	.530	.256	.127
40	2.70	2.42	2.02	1.68	1.30	.851	.681	.529	.255	.126
60	2.66	2.39	2.00	1.67	1.30	.848	.679	.527	.254	.126
120	2.62	2.36	1.98	1.66	1.29	.845	.677	.526	.254	.126
∞	2.58	2.33	1.96	1.645	1.28	.842	.674	.524	.253	.126

טבלת התפלגות F עבור $\alpha = 0.05$

df2/df1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	INF
1	161.448	199.5	215.707	224.583	230.162	233.99	236.768	238.883	240.543	241.882	243.91	245.95	248.013	249.052	250.095	251.143	252.196	253.253	254.314
2	18.5128	19	19.1643	19.2468	19.2964	19.33	19.3532	19.371	19.3848	19.3959	19.413	19.4291	19.4458	19.4541	19.4624	19.4707	19.4791	19.4874	19.4957
3	10.128	9.552	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.572	8.5494	8.5264
4	7.7086	6.944	6.5914	6.3882	6.2561	6.1631	6.0942	6.041	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.717	5.6877	5.6581	5.6281
5	6.6079	5.786	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.365
6	5.9874	5.143	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.099	4.06	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	5.5914	4.737	4.3468	4.1203	3.9715	3.866	3.787	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	5.3177	4.459	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	5.1174	4.257	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	4.9646	4.103	3.7083	3.478	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.913	2.845	2.774	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	4.8443	3.982	3.5874	3.3567	3.2039	3.0946	3.0123	2.948	2.8962	2.8536	2.7876	2.7186	2.6464	2.609	2.5705	2.5309	2.4901	2.448	2.4045
12	4.7472	3.885	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.341	2.2962
13	4.6672	3.806	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.671	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	4.6001	3.739	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5342	2.463	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	4.5431	3.682	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	4.494	3.634	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	4.4513	3.592	3.1968	2.9647	2.81	2.6987	2.6143	2.548	2.4943	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.104	2.0584	2.0107	1.9604
18	4.4139	3.555	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	4.3807	3.522	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.308	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.878
20	4.3512	3.493	3.0984	2.8661	2.7109	2.599	2.514	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	4.3248	3.467	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.366	2.321	2.2504	2.1757	2.096	2.054	2.0102	1.9645	1.9165	1.8657	1.8117
22	4.3009	3.443	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.938	1.8894	1.838	1.7831
23	4.2793	3.422	3.028	2.7955	2.64	2.5277	2.4422	2.3748	2.3201	2.2747	2.2036	2.1282	2.0476	2.005	1.9605	1.9139	1.8648	1.8128	1.757
24	4.2597	3.403	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9838	1.939	1.892	1.8424	1.7896	1.733
25	4.2417	3.385	2.9912	2.7587	2.603	2.4904	2.4047	2.3371	2.2821	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.711
26	4.2252	3.369	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1479	2.0716	1.9898	1.9464	1.901	1.8533	1.8027	1.7488	1.6906
27	4.21	3.354	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	4.196	3.34	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.236	2.19	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	4.183	3.328	2.934	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	4.1709	3.316	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	4.0847	3.232	2.8387	2.606	2.4495	2.3359	2.249	2.1802	2.124	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	4.0012	3.15	2.7581	2.5252	2.3683	2.2541	2.1665	2.097	2.0401	1.9926	1.9174	1.8364	1.748	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	3.9201	3.072	2.6802	2.4472	2.2899	2.175	2.0868	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.429	1.3519	1.2539
inf	3.8415	2.996	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.394	1.318	1.2214	1

טבלת F עבור $\alpha = 0.025$

120	60	40	30	24	20	15	12	10	9	8	7	6	5	4	3	2	$df_1=1$	/
1014	1010	1006	1001	997.2	993.1	984.9	976.7	968.6	963.3	956.7	948.2	937.1	921.8	899.6	864.2	799.5	647.8	$df_2=1$
39.49	39.48	39.47	39.47	39.46	39.45	39.43	39.41	39.4	39.39	39.37	39.36	39.33	39.3	39.25	39.17	39	38.51	2
13.95	13.99	14.04	14.08	14.12	14.17	14.25	14.34	14.42	14.47	14.54	14.62	14.73	14.88	15.1	15.44	16.04	17.44	3
8.309	8.36	8.411	8.461	8.511	8.56	8.657	8.751	8.844	8.905	8.98	9.074	9.197	9.365	9.605	9.979	10.65	12.22	4
6.069	6.123	6.175	6.227	6.278	6.329	6.428	6.525	6.619	6.681	6.757	6.853	6.978	7.146	7.388	7.764	8.434	10.01	5
4.904	4.959	5.012	5.065	5.117	5.168	5.269	5.366	5.461	5.523	5.6	5.696	5.82	5.988	6.227	6.599	7.26	8.813	6
4.199	4.254	4.309	4.362	4.415	4.467	4.568	4.666	4.761	4.823	4.899	4.995	5.119	5.285	5.523	5.89	6.542	8.073	7
3.728	3.784	3.84	3.894	3.947	4	4.101	4.2	4.295	4.357	4.433	4.529	4.652	4.817	5.053	5.416	6.06	7.571	8
3.392	3.449	3.505	3.56	3.614	3.667	3.769	3.868	3.964	4.026	4.102	4.197	4.32	4.484	4.718	5.078	5.715	7.209	9
3.14	3.198	3.255	3.311	3.365	3.419	3.522	3.621	3.717	3.779	3.855	3.95	4.072	4.236	4.468	4.826	5.456	6.937	10
2.944	3.004	3.061	3.118	3.173	3.226	3.33	3.43	3.526	3.588	3.664	3.759	3.881	4.044	4.275	4.63	5.256	6.724	11
2.787	2.848	2.906	2.963	3.019	3.073	3.177	3.277	3.374	3.436	3.512	3.607	3.728	3.891	4.121	4.474	5.096	6.554	12
2.659	2.72	2.78	2.837	2.893	2.948	3.053	3.153	3.25	3.312	3.388	3.483	3.604	3.767	3.996	4.347	4.965	6.414	13
2.552	2.614	2.674	2.732	2.789	2.844	2.949	3.05	3.147	3.209	3.285	3.38	3.501	3.663	3.892	4.242	4.857	6.298	14
2.461	2.524	2.585	2.644	2.701	2.756	2.862	2.963	3.06	3.123	3.199	3.293	3.415	3.576	3.804	4.153	4.765	6.2	15
2.383	2.447	2.509	2.568	2.625	2.681	2.788	2.889	2.986	3.049	3.125	3.219	3.341	3.502	3.729	4.077	4.687	6.115	16
2.315	2.38	2.442	2.502	2.56	2.616	2.723	2.825	2.922	2.985	3.061	3.156	3.277	3.438	3.665	4.011	4.619	6.042	17
2.256	2.321	2.384	2.445	2.503	2.559	2.667	2.769	2.866	2.929	3.005	3.1	3.221	3.382	3.608	3.954	4.56	5.978	18
2.203	2.27	2.333	2.394	2.452	2.509	2.617	2.72	2.817	2.88	2.956	3.051	3.172	3.333	3.559	3.903	4.508	5.922	19
2.156	2.223	2.287	2.349	2.408	2.465	2.573	2.676	2.774	2.837	2.913	3.007	3.128	3.289	3.515	3.859	4.461	5.872	20
2.114	2.182	2.246	2.308	2.368	2.425	2.534	2.637	2.735	2.798	2.874	2.969	3.09	3.25	3.475	3.819	4.42	5.827	21
2.076	2.145	2.21	2.272	2.332	2.389	2.498	2.602	2.7	2.763	2.839	2.934	3.055	3.215	3.44	3.783	4.383	5.786	22
2.041	2.111	2.176	2.239	2.299	2.357	2.467	2.57	2.668	2.731	2.808	2.902	3.023	3.184	3.408	3.751	4.349	5.75	23
2.01	2.08	2.146	2.209	2.269	2.327	2.437	2.541	2.64	2.703	2.779	2.874	2.995	3.155	3.379	3.721	4.319	5.717	24
1.981	2.052	2.118	2.182	2.242	2.301	2.411	2.515	2.614	2.677	2.753	2.848	2.969	3.129	3.353	3.694	4.291	5.686	25
1.954	2.026	2.093	2.157	2.217	2.276	2.387	2.491	2.59	2.653	2.729	2.824	2.945	3.105	3.329	3.67	4.266	5.659	26
1.93	2.002	2.069	2.133	2.195	2.253	2.364	2.469	2.568	2.631	2.707	2.802	2.923	3.083	3.307	3.647	4.242	5.633	27
1.907	1.98	2.048	2.112	2.174	2.232	2.344	2.448	2.547	2.611	2.687	2.782	2.903	3.063	3.286	3.626	4.221	5.61	28
1.886	1.959	2.028	2.092	2.154	2.213	2.325	2.43	2.529	2.592	2.669	2.763	2.884	3.044	3.267	3.607	4.201	5.588	29
1.866	1.94	2.009	2.074	2.136	2.195	2.307	2.412	2.511	2.575	2.651	2.746	2.867	3.027	3.25	3.589	4.182	5.568	30
1.724	1.803	1.875	1.943	2.007	2.068	2.182	2.288	2.388	2.452	2.529	2.624	2.744	2.904	3.126	3.463	4.051	5.424	40
1.581	1.667	1.744	1.815	1.882	1.945	2.061	2.169	2.27	2.334	2.412	2.507	2.627	2.786	3.008	3.343	3.925	5.286	60
1.433	1.53	1.614	1.69	1.76	1.825	1.945	2.055	2.157	2.222	2.299	2.395	2.515	2.674	2.894	3.227	3.805	5.152	120

בהצלחה!!!