

Лабораторная работа

«Регрессионный анализ»

Задание на лабораторную работу.

По результатам эксперимента построить линейную многофакторную модель и провести ее статистический анализ в MS Excel.

Ход работы:

1. Рассчитать множественный коэффициент корреляции, оценить его значимость, сделать выводы о виде связи.
2. С помощью метода наименьших квадратов получить точечные оценки параметров линейного многофакторного уравнения регрессии.
3. Оценить значимость полученных коэффициентов.
4. Рассчитать метрики качества для уравнения регрессии (MSE, RMSE, RSD). Оценить адекватность уравнения регрессии.
5. С помощью инструмента «Регрессия» надстройки «Анализ данных» в MS Excel постройте уравнение регрессии и опишите полученные результаты.
6. По полученным данным п.4, 5 сделать выводы относительно значимости коэффициентов уравнения, а также выводы относительно качества самого уравнения.

Таблица 6.1 - Варианты заданий. Вариант 1,3,5. (7, 9, 11)

X ₁ =C,%	X ₂ =Cr,%	X ₃ =Mn,%	X ₄ =Si,%	Y ₁ – относительное сужение, ψ , %	Y ₂ – относительное удлинение, δ , %	Y ₃ – предел прочности на разрыв, σ_b , МПа
0,43	0,11	0,71	1,71	50,75	13,82	1243
0,4	0,1	0,63	1,6	45,90	14,11	1258
0,43	0,12	0,7	1,72	50,19	13,64	1235
0,42	0,11	0,65	1,73	48,41	13,87	1247
0,43	0,13	0,67	1,62	51,23	14,12	1162
0,42	0,12	0,62	1,65	49,37	14,54	1190
0,45	0,12	0,69	1,63	55,67	15,35	1145
0,42	0,11	0,74	1,69	49,58	13,08	1269
0,45	0,11	0,71	1,62	55,03	15,27	1175
0,41	0,1	0,69	1,72	46,51	12,94	1298
0,42	0,11	0,69	1,66	48,54	13,51	1240
0,42	0,12	0,67	1,68	49,03	13,86	1231
0,43	0,11	0,66	1,78	50,98	14,38	1234
0,42	0,09	0,7	1,64	49,34	14,03	1267
0,42	0,12	0,68	1,67	49,04	13,52	1217
0,43	0,09	0,7	1,59	51,52	14,81	1230
0,42	0,1	0,72	1,65	48,95	13,48	1264
0,44	0,14	0,71	1,69	53,66	13,94	1156
0,44	0,12	0,65	1,61	52,59	15,33	1163
0,44	0,14	0,67	1,62	52,71	14,44	1147
0,44	0,1	0,69	1,63	52,27	14,9	1208
0,41	0,11	0,75	1,6	47,89	12,76	1268
0,41	0,09	0,66	1,67	46,88	13,9	1283
0,42	0,12	0,64	1,61	48,29	13,99	1201
0,43	0,1	0,69	1,63	50,54	14,48	1235
0,45	0,11	0,67	1,69	55,48	15,74	1164
0,4	0,11	0,65	1,6	45,65	13,37	1251
0,43	0,11	0,7	1,68	50,95	14,06	1229
0,42	0,11	0,66	1,68	49,09	14,15	1243
0,42	0,08	0,64	1,61	49,91	15,42	1246
0,42	0,12	0,71	1,63	48,29	12,92	1231
0,43	0,11	0,73	1,63	51,53	13,82	1223
0,42	0,11	0,65	1,63	49,50	14,54	1221
0,42	0,13	0,72	1,59	48,29	12,78	1223
0,44	0,12	0,69	1,64	53,62	14,7	1171
0,42	0,11	0,67	1,58	49,88	14,49	1209
0,41	0,1	0,65	1,64	47,61	14,02	1246
0,41	0,12	0,68	1,69	47,90	13,15	1245
0,42	0,12	0,74	1,69	49,49	12,82	1250
0,43	0,12	0,69	1,58	51,23	14,27	1183
0,43	0,09	0,67	1,67	50,47	14,71	1246
0,42	0,11	0,65	1,68	49,27	14,27	1226
0,43	0,1	0,69	1,59	50,60	14,51	1219
0,45	0,13	0,67	1,68	53,94	14,79	1152

Продолжение таблицы 6.1

0,43	0,11	0,64	1,62	50,01	14,66	1195
0,43	0,13	0,71	1,71	50,11	13,22	1229
0,41	0,12	0,67	1,75	46,36	12,74	1269
0,42	0,09	0,67	1,72	48,51	14,09	1287
0,43	0,12	0,67	1,64	51,32	14,39	1180
0,42	0,11	0,63	1,71	49,97	14,74	1218

Таблица 6.2 - Варианты заданий. **Вариант 2,4,6. (8,10,12)**

X ₁ =C,%	X ₂ =Al,%	X ₃ =Mn,%	X ₄ =Si,%	Y1 – предел текучности, σ_T , МПа	Y2 – относительное удлинение, δ , %	Y3 – предел прочности на разрыв, σ_B , МПа
0,20	0,03	1,52	0,20	971	13,9	1037
0,21	0,04	1,44	0,23	925	15,3	993
0,19	0,03	1,46	0,26	931	16,7	996
0,20	0,03	1,40	0,21	881	14,5	948
0,20	0,03	1,48	0,25	958	16,3	1025
0,21	0,04	1,51	0,21	977	14,4	1046
0,21	0,03	1,56	0,27	1039	17,1	1107
0,20	0,04	1,43	0,27	923	17,4	991
0,20	0,04	1,47	0,27	958	17,4	1026
0,19	0,05	1,42	0,25	897	16,4	961
0,22	0,02	1,44	0,23	943	15,2	1016
0,21	0,03	1,48	0,28	977	17,7	1046
0,20	0,04	1,43	0,21	908	14,4	976
0,20	0,04	1,43	0,30	941	19,2	1009
0,20	0,04	1,38	0,30	905	19,3	974
0,19	0,04	1,40	0,23	876	15,7	942
0,19	0,03	1,47	0,18	905	13,2	968
0,21	0,03	1,47	0,31	984	19,3	1055
0,20	0,03	1,47	0,32	979	19,5	1046
0,21	0,03	1,41	0,27	921	17,4	991
0,18	0,03	1,48	0,38	986	22,6	1049
0,19	0,04	1,43	0,25	906	16,4	970
0,19	0,04	1,48	0,25	940	16,4	1003
0,20	0,04	1,44	0,34	969	20,6	1038
0,21	0,03	1,41	0,29	927	18,5	998
0,20	0,03	1,46	0,29	959	18,5	1027
0,20	0,04	1,43	0,25	921	16,4	989
0,21	0,04	1,45	0,35	990	21,3	1063
0,20	0,04	1,45	0,29	953	18,2	1021
0,21	0,04	1,46	0,33	985	20,4	1056
0,20	0,04	1,42	0,22	899	14,9	966
0,21	0,03	1,47	0,33	993	20,3	1063
0,20	0,03	1,49	0,30	984	18,6	1051
0,20	0,02	1,47	0,17	917	12,5	983
0,22	0,05	1,50	0,28	1010	17,5	1082

Продолжение таблицы 6.2

0,21	0,02	1,46	0,35	1005	21,3	1077
0,22	0,03	1,42	0,30	955	19	1028
0,20	0,03	1,47	0,26	955	17	1022
0,21	0,03	1,49	0,34	1017	20,4	1089
0,21	0,04	1,40	0,35	941	21,2	1011
0,20	0,04	1,43	0,32	948	19,8	1016
0,22	0,04	1,44	0,16	909	12,2	980
0,20	0,03	1,47	0,22	941	14,9	1008
0,22	0,04	1,44	0,22	935	14,8	1007
0,20	0,05	1,47	0,28	964	17,9	1032
0,21	0,04	1,44	0,27	948	17,2	1019
0,22	0,03	1,49	0,26	985	16,6	1056
0,21	0,03	1,43	0,34	972	20,7	1043
0,20	0,03	1,43	0,35	949	21,3	1017
0,21	0,03	1,46	0,23	943	15,5	1011

Таблица 6.3 – Испытания на параллельных опытах. **Вариант 1,3,5. (7, 9, 11)**

X ₁ =C,%	X ₂ =Al,%	X ₃ =Mn,%	X ₄ =Si,%	y1	y2	y3
0,41	0,1	0,63	1,63	34,56	12,88	1123
0,41	0,1	0,63	1,63	34,24	13,35	1113
0,41	0,1	0,63	1,63	34,98	13,55	1112
0,41	0,1	0,63	1,63	34,44	13,02	1121
0,41	0,1	0,63	1,63	35,67	13,19	1123
0,41	0,1	0,63	1,63	35,10	13,10	1107
0,41	0,1	0,63	1,63	34,90	13,14	1112

Таблица 6.4 – Испытания на параллельных опытах. **Вариант 2,4,6. (8,10,12)**

X ₁ =C,%	X ₂ =Al,%	X ₃ =Mn,%	X ₄ =Si,%	y1	y2	y3
0,41	0,1	0,63	1,63	908	16,02	973
0,41	0,1	0,63	1,63	912	15,61	972
0,41	0,1	0,63	1,63	906	15,65	964
0,41	0,1	0,63	1,63	908	15,85	963
0,41	0,1	0,63	1,63	907	15,97	965
0,41	0,1	0,63	1,63	903	16,00	970
0,41	0,1	0,63	1,63	905	16,06	970

Замечание к п.5.

В качестве одной из результирующих таблиц «Регрессионный анализ» надстройки «Анализ данных» является таблица «Дисперсионный анализ», в которой представлен анализ значимости регрессии с помощью оценки коэффициента детерминации. В строке «Регрессия» представлены кол-во степеней свободы, сумма квадратов отклонения и дисперсия, объясненные регрессией, а в строке «Остаток» те же показатели относительно остатков. В

случае линейной регрессии коэффициент детерминации может быть рассчитан по упрощенной формуле:

$$R^2 = \frac{\sum_{i=1}^n (y(x_i) - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{SS_{\text{регр}}^2}{SS_{\text{общ}}^2},$$

т.к. в случае линейной регрессии выполняется равенство:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (y(x_i) - \bar{y})^2 + \sum_{i=1}^n (y(x_i) - y_i)^2.$$

Значимость коэффициента детерминации оценивается по критерию Фишера (см. лекцию).

Контрольные вопросы

1. Как вы понимаете понятие значимости коэффициентов регрессии?
2. Сформулировать понятие адекватности уравнения регрессии?
3. Как выбрать наилучшее уравнение регрессии?
4. Каким требованиям должны подчиняться регрессионные остатки?
5. Способы линеаризации регрессионной модели относительно коэффициентов.
6. Для чего проводятся параллельные опыты?
7. Множественный коэффициент корреляции, его свойства.