

- 1.(25점) 화학실험에서 반응온도( $X_1$ ), 압력( $X_2$ ), 촉매량( $X_3$ )이 반응속도( $Y$ )에 미치는 영향의 파악을 [20분] 위해 다중선형회귀모형( $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \epsilon_i$ )을 적합한 결과이다.

Variable	Estimate	Standard Error	t-value	p-value
Intercept	74.8155	76.6245	0.9764	0.3379
X1	1.1132	1.1398	0.9771	0.3375
X2	1.8106	0.3790	4.7778	0.0001
X3	-0.4824	0.3486	-1.3841	0.1781

- 가. 현재까지의 분석결과로 볼 때 반응변수에 영향을 주는 설명변수는 무엇인가?  
(어떤 설명변수가 포함되나? 예: ( $X_1$ ), ( $X_2$ ), ( $X_1, X_2$ ), ( $X_1, X_3$ ), ..., 답이 여러개 일 수도 있음)
- 나. 최종적인 모형에 대한 결론을 내리기 전에 어떠한 분석 과정이 추가적으로 필요한가?  
(어떤 과정 또는 어떤 적합이 필요한가를 구체적으로 밝힐 것)
- 다. 추정값  $\hat{\beta}_2 = 1.81$ 의 의미를 적어라. (설명변수, 반응변수, 기술기 용어는 사용하지 말 것)
- 라. (1) 설명변수  $X_2$ 의 추가변수그림에서 가로축과 세로축에 어떤 잔차를 그려야 하는가?  
(2)  $X_2$ 의 추가변수그림의 형태에 대하여 기술하라.

- 2.(25점) 폐활량( $Y$ )에 영향을 주는 원인을 알아보기 위해 13명을 대상으로 폐의 크기( $X_1$ ), 신장( $X_2$ ), [20분] 체중( $X_3$ ), 나이( $X_4$ ), 혈압( $Y$ )을 조사하고 다중회귀모형  $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \epsilon_i$ 을 적합한 결과와 분산분석표, 순차분산분석표이다.

Variable	Est.	SE	p-value	Source	SS	Source	Sq.SS	Source	Sq.SS
intercept	-619.5	124.1	0.004	Regression	2536	x1	24.2	x4	1084.5
x1	-1.38	.680	0.076	Residual	231	x2	2426.0	x3	28.3
x2	5.37	.728	0.000	Total	2767	x3	41.8	x2	1315.2
x3	-4.23	1.41	0.775			x4	44.0	x1	108.0
x4	-1.21	9.20	0.226						

- 가.  $H_0: \beta_3 = 0$  대  $H_1: \beta_3 \neq 0$  검정의 결과로 의사는 「체중( $X_3$ )은 폐활량에 영향을 주지 않는다」라는 결론을 내렸다. 이것이 적절한 결론인가? 그 이유를 50자 이내로 구체적으로 설명하라.
- 나. 첫 번째 설명변수 폐의 크기( $X_1$ )에 대한 편상관계수를 구하라.
- 다. 의사는 폐의 크기( $X_1$ )가 크면 폐활량( $Y$ )도 많아지는데  $X_1$ 에 대한  $\hat{\beta}_1$ 이 음수가 나온 것을 이해하지 못하고 있다. 의사가 납득할 수 있도록 50자 이내로 구체적으로 설명하라.
- 라. (1)  $H_0: \beta_3 = \beta_4 = 0$  대  $H_1: \text{not } H_0$  을 검정하라, [참고:  $F(.05; df_1, df_2) = 4.46$ ]  
(2) 검정 결과를 구체적으로 설명하라. (귀무가설을 기각 / 기각하지 못함은 결과가 아님)

3.(20점) 음료 제품의 선호도( $Y$ )와 수분의 함량( $X_1$ ), 당도( $X_2$ ), 투명도( $X_3$ )의 관계를 알아보기 위해

[15분] 다중선형회귀모형  $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \epsilon_i$ 을 생각하였다. 자료와 주어진 통계값을

이용하여 적절한 모형인지 검정하라. (가설, 검정통계량과 분포, 결론을 구체적으로 적어라.)

[참고:  $SSR = 30.97$ ,  $SSE = 51.20$ ,  $SSPE = 4.96$ ,  $F(.05; df_1, df_2) = 4.70$ ]

번호	선호도( $Y$ )	수분( $X_1$ )	당도( $X_2$ )	투명도( $X_3$ )	번호	선호도( $Y$ )	수분( $X_1$ )	당도( $X_2$ )	투명도( $X_3$ )
1	6.6	225	46	0.5	11	6.9	255	39.9	1.1
2	6.9	285	46	0.5	12	6.3	255	70.1	1.1
3	7.9	225	64	0.5	13	4.0	255	55	0.09
4	6.1	285	64	0.5	14	8.6	255	55	2.11
5	9.2	225	46	1.7	15	10.1	255	55	1.1
6	6.8	285	46	1.7	16	9.9	255	55	1.1
7	10.4	225	64	1.7	17	12.2	255	55	1.1
8	7.3	285	64	1.7	18	9.7	255	55	1.1
9	9.8	204.5	55	1.1	19	9.7	255	55	1.1
10	5.0	305.5	55	1.1	20	9.6	255	55	1.1

4.(10점) 환자만족도( $Y$ )와 나이( $X_1$ ), 심각성( $X_2$ ),불안도( $X_3$ )의 관계를 알아보기 위해 환자 20명을 대상

[10분] 으로 적합결과 추정회귀식  $\hat{y} = 3.5 - 1.4x_1 + 8.9x_2 + 6.7x_3$ 를 얻었다.  $X_1 = 32$ ,  $X_2 = 8$ ,  $X_3 = 3$ 인

환자만족도 예측값의 95% 신뢰구간을 구하라. [참고:  $h_{00} = 1.35$ ,  $SSE = 3.30$ ,  $t(.025; df) = 2.12$ ]