

단순선형회귀분석4

Park Beomjin¹

¹University of Seoul

1 예제1 : 물류 배송 데이터

- 물류 창고는 판매 센터에 다양한 제품을 제공한다..
- 이러한 품목의 배송이 지연되면 추가 비용이 발생하고 정시에 배송이 된다면 추가적인 비용이 발생하지 않는다.
- 우리는 배송이 지연됨에 따라 발생하는 추가비용을 알아보기 위해 회귀모형을 적합하기로 하였다.
- 절편이 존재하는 회귀모형과 절편이 존재하지 않는 회귀모형을 각각 적합하고 차이점을 확인하여라.
- 데이터 읽기

```
DATA reg.logistics;  
input delay cost @@;  
cards;  
1 3.47  
2 1.02  
2 4.32  
3 5.03  
4 9.27  
6 14.29  
7 12.30  
8 18.25  
9 17.95  
9 19.81  
;  
RUN;
```

- 두 변수간의 산점도를 그려보고 어떠한 관계가 있을지 예상하여 보자.

```
PROC SGPLOT data = reg.logistics;
scatter X = delay Y = cost;
RUN;
```

- 반응변수(Y)를 cost 설명변수(X)를 delay로 하는 회귀모형을 고려하고 절편이 있는 회귀모형과 없는 회귀모형을 각각 적합해보라.
- 절편이 존재하는 회귀모형 (M1)과 절편이 존재하지 않는 회귀모형 (M2) 적합

```
PROC REG data = reg.logistics;
M1 : model cost = delay;
M2 : model cost = delay/noint;
RUN;
```

- M1과 M2의 회귀모형 적합 결과를 비교하여 보자. 어떤 모형이 더 잘 적합되었다고 볼 수 있는가?

2 예제2 : 주교재 예제 2.27

- 예제 2.27 데이터는 화학 공정에서 liquid의 양에 따른 특정 성분의 흡수량에 대한 데이터이다.
- 변수 Y1은 liquid에 따른 흡수된 성분의 양을 나타내고 변수 Y2는 liquid의 양 대비 흡수된 퍼센트를 나타낸다.
- Q1 : liquid와 Y1 그리고 Y2의 산점도를 그려보고 변수들의 관계를 확인하여보자.
- Q2 : liquid를 설명변수(X)로 하고 Y1과 Y2를 반응변수로 하는 회귀모형을 각각 적합하고 결과를 확인하여라.
- Q3 : 어느 모형이 특정 성분 흡수에 대한 예측을 더 잘할 것으로 생각되는가?
- Q4 : 각 회귀모형을 통해 liquid가 800, 1160, 1490에 대한 예측값을 산출하고 비교하여라.

```
PROC IMPORT datafile = "C:\Users\User\Desktop\19 regression class\Data\Week5 Data"
dbms = dlm replace out = reg.absorption;
getnames = yes;
delimiter = ",";
RUN;
```

```
/* SAS CODE for Q1 */
PROC SGPLOT data = reg.absorption;
scatter X = Liquid Y = Y1;
RUN;
```

```
PROC SGPLOT data = reg.absorption;
scatter X = Liquid Y = Y2;
RUN;
```

```
/* SAS CODE for Q2 */
PROC REG data = reg.absorption outest = reg.reg_out;
M1 : model Y1 = Liquid;
M2 : model Y2 = Liquid;
M3 : model Y1 = Liquid/noint;
M4 : model Y2 = Liquid/noint;
RUN;
```

```
/* SAS CODE for Q4 */
DATA reg.test_absorption;
input Liquid Y1 Y2 @@@;
cards;
800 43.8 5.48
1160 79.6 6.86
1490 98.6 6.62
;
RUN;
```

```
PROC SCORE data = reg.test_absorption score = reg.reg_out out = reg.predicted_abs
var Liquid;
```

```
RUN;
```

```
/* 특정 모형으로만 예측 값을 산출 할 수 있음
```

```
PROC SCORE data = reg.test_absorption score = reg.reg_out(where = (_MODEL_ = "M1
```

```
out = reg.predicted_abs type = parms;
```

```
var Liquid;
```

```
RUN;
```

```
*/
```

```
DATA reg.rpredicted_abs;
```

```
set reg.predicted_abs;
```

```
m1Y1 = M1 / Liquid * 100;
```

```
m3Y1 = M3 / Liquid * 100;
```

```
RUN;
```