

퀴즈 : 6주차 (이름 :)

1. 다음 중 표준회귀모형에 대한 설명으로 맞지 않는 것은? ()

- ① 관측된 변수의 측정단위에 관계없이 표준회귀모형을 이용해야 한다.
- ② 독립변수의 상대적 중요도를 알고자 할 경우 표준회귀모형을 이용한다.
- ③ 추정된 회귀계수는 독립변수 표준편차가 한 단위가 변할 때 종속변수의 표준편차 크기의 변화를 나타낸다.
- ④ 표준회귀모형으로 추정된 베타회귀계수와 표준화되지 않은 회귀모형으로 추정된 회귀계수 사이에 일정한 관계가 존재한다.

2. 다음 다중회귀모형 추정결과에 대한 설명으로 맞지 않는 것은? 단, sales는 매출액을 promotion은 홍보비 지출액을 그리고 research는 연구개발 지출액을 각각 나타낸다. ()

- ① promotion의 표준편차는 0.81649이고, research의 표준편차는 0.5773이다.
- ② promotion의 표준편차는 0.5이고, research의 표준편차는 0.707107이다.
- ③ 표준회귀모형은 상수항을 추정할 필요가 없다.
- ④ 종속변수인 sales의 표준편차는 0.957427이다.

```
call:
lm(formula = sales ~ promotion + research)
```

```
Residuals:
    1      2      3      4
-0.25  0.25 -0.25  0.25
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -4.2500     1.9203  -2.213   0.270
promotion      1.5000     0.5000   3.000   0.205
research       2.0000     0.7071   2.828   0.216
```

```
Residual standard error: 0.5 on 1 degrees of freedom
Multiple R-squared:  0.9091,    Adjusted R-squared:  0.7273
F-statistic:    5 on 2 and 1 DF,  p-value: 0.3015
```

```
call:
lm(formula = tsales ~ tpromotion + tresearch)
```

```
Residuals:
    1      2      3      4
-0.2611  0.2611 -0.2611  0.2611
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.695e-17  2.611e-01   0.000   1.000
tpromotion    1.279e+00  4.264e-01   3.000   0.205
tresearch     1.206e+00  4.264e-01   2.828   0.216
```

```
Residual standard error: 0.5222 on 1 degrees of freedom
Multiple R-squared:  0.9091,    Adjusted R-squared:  0.7273
F-statistic:    5 on 2 and 1 DF,  p-value: 0.3015
```

3. 아래에 주어진 단순회귀모형 추정결과를 이용하여 다음의 다중회귀모형의 추정식을 바르게 나낸 것은? 단, 변수에 대한 설명은 2번과 동일하다. ()

$$\widehat{sales} = \widehat{\beta}_1 + \widehat{\beta}_2 promotion + \widehat{\beta}_3 research$$

- ① $\widehat{\beta}_1 = 0.5, \widehat{\beta}_2 = 2.0$
- ② $\widehat{\beta}_1 = 1.5, \widehat{\beta}_2 = 0.5$
- ③ $\widehat{\beta}_1 = 2.0, \widehat{\beta}_2 = 1.5$
- ④ $\widehat{\beta}_1 = 1.5, \widehat{\beta}_2 = 2.0$

4. 다음 중 Gauss-Markov 정리에 대한 설명으로 맞지 않는 것은? ()

- ① 추정방법은 최소자승법이어야 한다. ② 모든 회귀모형에서 성립한다.
- ③ 선형추정량이어야 한다. ④ 불편추정량이어야 한다.

5. OLS 추정량이 선형 추정량인 이유는? ()

- ① 설명변수가 확정변수라는 가정 때문이다
- ② 교란항의 평균이 0이라는 가정 때문이다
- ③ 교란항의 분산이 동분산이라는 가정 때문이다
- ④ 교란항과 설명변수는 서로 독립이라는 가정 때문이다.

6. OLS 추정량이 불편 추정량인 이유는? ()

- ① 설명변수가 확정변수라는 가정 때문이다
- ② 설명변수가 확정변수, 교란항의 평균이 0이라는 가정 때문이다
- ③ 설명변수가 확정변수, 교란항의 분산이 동분산이라는 가정 때문이다
- ④ 교란항과 설명변수는 서로 독립이고, 교란항의 분산은 동분산이라는 가정 때문이다

```
> olspromotion<-lm(promotion~research)
> respromotion<-resid(olspromotion)
> ols2<-lm(sales~respromotion+research)
> summary(ols2)
```

```
call:
lm(formula = sales ~ respromotion + research)

Residuals:
    1     2     3     4 
-0.25  0.25 -0.25  0.25 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.0000     0.7906   1.265   0.426
respromotion   1.5000     0.5000   3.000   0.205
research       0.5000     0.5000   1.000   0.500

Residual standard error: 0.5 on 1 degrees of freedom
Multiple R-squared:  0.9091,    Adjusted R-squared:  0.7273 
F-statistic:    5 on 2 and 1 DF,  p-value: 0.3015
```

```
> olsresearch<-lm(research~promotion)
> resresearch<-resid(olsresearch)
> ols3<-lm(sales~promotion+resresearch)
> summary(ols3)
```

```
call:
lm(formula = sales ~ promotion + resresearch)

Residuals:
    1     2     3     4 
-0.25  0.25 -0.25  0.25 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.7500     0.7500   1.000   0.500
promotion     0.5000     0.3536   1.414   0.392
resresearch    2.0000     0.7071   2.828   0.216

Residual standard error: 0.5 on 1 degrees of freedom
Multiple R-squared:  0.9091,    Adjusted R-squared:  0.7273 
F-statistic:    5 on 2 and 1 DF,  p-value: 0.3015
```

7. 다음 다중회귀모형의 추정결과에 대한 설명 중 맞지 않는 것은? 단, X는 설명변수의 데이터 행렬이다.()

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

- ① $\hat{\beta}_3$ 과 $\hat{\beta}_2$ 의 공분산은 0.25이다
- ② $\hat{\beta}_2$ 과 $\hat{\beta}_3$ 의 공분산은 0.25이다
- ③ $\hat{\beta}_2$ 의 표준오차는 0.25이다.
- ④ 교란항의 분산은 0.25이다.

```

> x
      [,1] [,2] [,3]
[1,]    1    1    2
[2,]    1    2    1
[3,]    1    3    1
[4,]    1    2    2
> xpx=t(x)%*%x
> xpx
      [,1] [,2] [,3]
[1,]    4    8    6
[2,]    8   18   11
[3,]    6   11   10
> xpxinv=solve(xpx)
> xpxinv
      [,1] [,2] [,3]
[1,] 14.75 -3.5  -5
[2,] -3.50  1.0   1
[3,] -5.00  1.0   2

> varcov<-0.25*xpxinv
> varcov
      [,1] [,2] [,3]
[1,]  3.6875 -0.875 -1.25
[2,] -0.8750  0.250  0.25
[3,] -1.2500  0.250  0.50

```

8. 다음 중 결정계수에 대한 설명으로 맞지 않는 것은? ()

- ① 다중회귀분석에서는 결정계수가 클수록 모형의 설명력은 높다
- ② 종속변수의 형태와 관계없이 두 모형의 결정계수를 비교할 수 있다
- ③ 다중회귀분석에서 결정계수 계산식은 종속변수의 평균이 0인 경우와 그렇지 않은 경우에 따라 다르다
- ④ 다중회귀분석에서 통계적으로 유의하지 않은 설명변수가 추가되면 결정계수의 값은 변화가 없다

9. 다음 중 분산 추정에 대한 설명으로 맞지 않는 것은? ()

- ① 다중회귀분석에서 교란항의 분산은 잔차의 제곱의 합을 관측치의 수에서 2를 뺀 값으로 나누어 구한다.
- ② 회귀계수의 분산을 추정하기 위해서는 교란항의 분산 추정을 먼저 해야 한다
- ③ 회귀계수에 대한 가설검정을 위해서 분산 추정이 필요하다,
- ④ 교란항의 분산이 커지면 모든 회귀계수의 분산이 커진다

퀴즈 : 7주차 (이름 :)

1. 다음 중 가설검정에 대한 설명으로 맞지 않는 것은? ()

- ① 단일 검정은 t-검정만 가능한데 자유도는 자료의 수에서 상수항을 포함한 독립변수의 수(즉, $n-k$)를 뺀 것이다.
- ② 결합 검정은 F-검정을 이용하는데 분자의 자유도는 귀무가설의 수(즉, 제약식의 수)이다.
- ③ 전체 검정은 F-검정을 이용하는데 분자의 자유도는 상수항을 제외한 독립변수의 수(즉, $k-1$)이다.
- ④ 단순회귀분석에서는 단일 검정과 전체 검정의 결과가 항상 동일하다.

☞ 다음 다중회귀모형 추정결과를 보고 물음에 답하라. 단, sales는 매출액(단위 : 10억 원), promotion은 홍보비 지출액(단위 : 천만 원), research는 연구개발 지출액(단위 : 1억 원)을 나타내며, tsales, tpromotion, tresearch는 각각 매출액, 홍보비 지출액, 연구개발 지출액을 표준화한 변수를 나타낸다. 한편, lsales, lpromotion, lsearch는 각각 매출액, 홍보비 지출액, 연구개발 지출액에 로그를 취한 변수를 나타내고, nsales는 sales에 10을 곱한 변수이다.

2. 다음 중 경제적 유의성과 통계적 유의성(5% 유의수준)을 모두 가진 변수는? ()

- ① promotion ② tpromotion ③ lpromotion ④ lsales

3. sales, promotion, research을 이용한 다중회귀모형 추정결과를 설명한 것 중 맞지 않는 것은? 단, 통계적 유의성 여부는 5% 유의수준에서 하라 ()

- ① 홍보비 지출액은 모형에서 제외되어야 한다
- ② 연구개발 지출액은 모형에서 제외되어야 한다
- ③ 홍보비 지출액 및 연구개발 지출액은 모두 모형에서 제외되어야 한다
- ④ 개별 독립변수에 대한 t-검정의 임계값은 12.706이고, F-검정의 임계값은 161.45이다

```
call:
lm(formula = nsales ~ promotion + research)

Residuals:
    1     2     3     4 
-2.5  2.5 -2.5  2.5 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -42.500     19.203  -2.213   0.270
promotion      15.000      5.000   3.000   0.205
research      20.000      7.071   2.828   0.216

Residual standard error: 5 on 1 degrees of freedom
Multiple R-squared:  0.9091,    Adjusted R-squared:  0.7273 
F-statistic:    5 on 2 and 1 DF,  p-value: 0.3015

call:
lm(formula = tsales ~ tpromotion + tresearch - 1)

Residuals:
    1     2     3     4 
-0.2611  0.2611 -0.2611  0.2611 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
tpromotion   1.2792     0.3015   4.243  0.0513 .
tresearch    1.2060     0.3015   4.000  0.0572 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3693 on 2 degrees of freedom
Multiple R-squared:  0.9091,    Adjusted R-squared:  0.8182 
F-statistic:   10 on 2 and 2 DF,  p-value: 0.09091
```

```
call:
lm(formula = sales ~ promotion + research)

Residuals:
    1     2     3     4 
-0.25  0.25 -0.25  0.25 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -4.2500     1.9203  -2.213   0.270
promotion      1.5000     0.5000   3.000   0.205
research      2.0000     0.7071   2.828   0.216

Residual standard error: 0.5 on 1 degrees of freedom
Multiple R-squared:  0.9091,    Adjusted R-squared:  0.7273 
F-statistic:    5 on 2 and 1 DF,  p-value: 0.3015
```

```
call:
lm(formula = lsales ~ lpromotion + lresearch)

Residuals:
    1     2     3     4 
0.01100 -0.01881  0.01881 -0.01100 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.10181    0.05329  -20.68  0.0308 *
lpromotion   1.61672    0.05428   29.78  0.0214 *
lresearch    1.57370    0.06187   25.44  0.0250 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03082 on 1 degrees of freedom
Multiple R-squared:  0.9989,    Adjusted R-squared:  0.9968 
F-statistic: 465.2 on 2 and 1 DF,  p-value: 0.03277
```

☞ 다음 다중회귀모형 추정결과를 보고 물음에 답하라. 단, sales는 매출액, promotion은 홍보비 지출액, research는 연구개발 지출액, exchange는 달러 환율을 나타낸다.

4. 결정계수의 값은 얼마인가? ()

5. F-statistic : 1.382 on (1) 및 (2)의 값은 각각 얼마인가?

(1) ()

(2) ()

6. research와 exchange가 동시에 sales에 영향을 주지 않는다는 귀무가설을 검정한 내용에 대한 설명 중 맞지 않는 것은? ()

① 5% 유의수준에서 귀무가설을 허용한다.

② F-통계량의 값은 12.25이다

③ F-통계량의 자유도는 분자의 자유도가 2, 분모의 자유도가 1이다.

④ 제약이 가해지지 않은 모형의 R^2 는 제약이 가해진 모형의 R^2 보다 항상 크다.

```
call:
lm(formula = sales ~ promotion + research + exchange)
```

```
Residuals:
    1      2      3      4      5
0.05978 -0.59783  0.41848  0.29891 -0.17935
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.8424    1.8744  -0.983   0.505
promotion      0.5761    0.3966   1.453   0.384
research      0.5109    0.5212   0.980   0.506
exchange      0.5924    0.4745   1.248   0.430
```

```
Residual standard error: 0.8109 on 1 degrees of freedom
Multiple R-squared:  0.8356,    Adjusted R-squared:  0.3424
F-statistic: 1.694 on 3 and 1 DF,  p-value: 0.5017
```

```
call:
lm(formula = sales ~ promotion)
```

```
Residuals:
    1      2      3      4      5
-0.19231 -0.76923 -0.34615  1.23077  0.07692
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.6154    0.9946   0.619   0.580
promotion      0.5769    0.3814   1.513   0.228
```

```
Residual standard error: 0.8697 on 3 degrees of freedom
Multiple R-squared:  0.4327,    Adjusted R-squared:  0.2436
F-statistic: 2.288 on 1 and 3 DF,  p-value: 0.2276
```

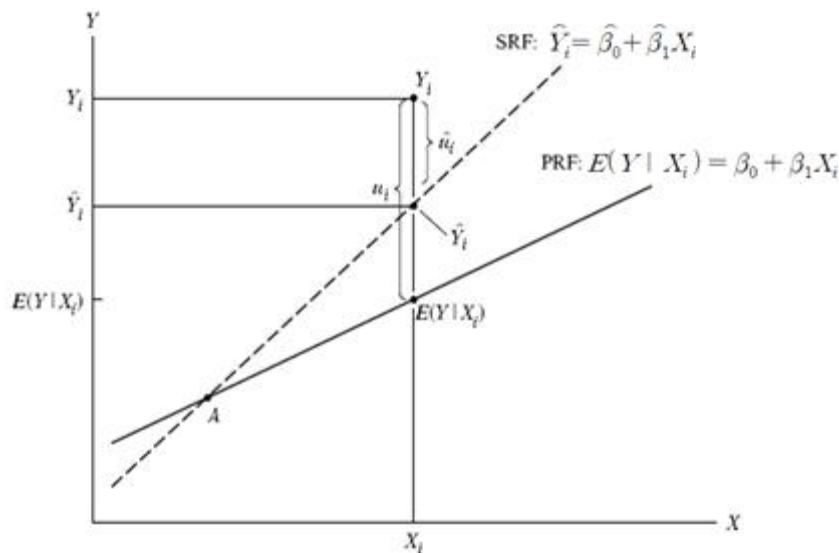
7. 다음 그림을 보고 각 기호에 해당하는 적절한 용어를 써 넣어라

① \hat{u}_i ()

② u_i ()

③ \hat{Y}_i ()

④ $E(Y|X_i)$ ()



☞ 다음 설명의 참, 거짓을 판별하라.

8. 예측오차분산은 개별예측의 경우가 평균예측의 경우보다 큰데 이 사실은 점예측 및 구간예측에 모두 영향을 준다 ()

- ① 거짓
- ② 참

9. 설명변수 X 의 모든 값에 대해 개별예측의 예측구간이 평균예측의 예측구간보다 넓고, X 의 값이 평균으로부터 멀어질수록 예측구간의 폭이 작아진다.()

- ① 거짓
- ② 참