

SAS 실습과제 4

Correlation analysis / Logistic regression

1. 중간고사 6.(1) 문제에서 작성한 work.midterm 데이터셋을 이용하여 혈중 카드뮴 농도와 TSH, fT4, TPOAb 와의 Correlation analysis 를 실시하고 다음의 Table 3 을 완성하시오.
(회색 글씨는 표시 형식을 보여주는 예이므로 삭제하고, 실제 숫자를 확인후 검은색 글씨로 채워 넣으시오)

Table 3. Correlations between blood cadmium levels and thyroid hormone levels

	Pearson's correlation coefficients (<i>P</i> value)	
	Men (n=1055)	Women (n=896)
TSH, μ IU/mL	-0.0582 (0.0585)	0.085 (0.0104)
fT4, ng/dL	-0.117 (0.0001)	0.013 (0.692)
TPOAb, IU/mL	0.021 (0.4826)	0.0301 (0.3671)

TSH, thyroid-stimulating hormone; fT4, free thyroxine; TPOAb, anti-thyroid peroxidase antibody.

[SAS code] PROC SORT DATA=work.midterm;
BY sex;
RUN;
PROC CORR DATA=work.midterm;
VAR HE_Cd HE_TSH HE_fT4 HE_TPOAb;
BY sex; /* 성별 변수로 그룹화 */
CORR;
RUN;

2. 중간고사 6.(1) 문제에서 작성한 work.midterm 데이터셋을 이용하여, 갑상선기능저하증 상태를 나타내는 ‘hypothyroidism’ 변수를 다음의 조건을 이용해서 새롭게 생성하여 추가한 뒤 work.quiz4 라는 데이터셋으로 저장하는 SAS code 를 작성하시오.

* 다음의 1) 또는 2) 또는 3) 의 조건 중 하나를 만족하는 경우 hypothyroidism=1

그렇지 않은 경우 [다음 1),2),3)의 조건 중 어느 하나도 만족하지 않는 경우] hypothyroidism=0

1) subclinical hypothyroidism

- $4 \mu\text{IU/mL} < \text{TSH}$ and $\text{TSH} \leq 10 \mu\text{IU/mL}$ and $0.7 \text{ ng/dL} \leq \text{fT4}$ and $\text{fT4} \leq 2.1 \text{ ng/dL}$

2) overt hypothyroidism

- $4 \mu\text{IU/mL} < \text{TSH}$ and $\text{TSH} \leq 10 \mu\text{IU/mL}$ and $\text{fT4} < 0.7 \text{ ng/dL}$

3) overt hypothyroidism

- $\text{TSH} > 10 \mu\text{IU/mL}$

[SAS code]

```
data work.quiz4;
  set work.midterm;
  if (4 < HE_TSH <= 10 and 0.7 <= HE_fT4 <= 2.1) then hypothyroidism = 1;
  else if (4 < HE_TSH <= 10 and HE_fT4 < 0.7) then hypothyroidism = 1;
  else if (HE_TSH > 10) then hypothyroidism = 1;
  else hypothyroidism = 0;
run;
```

3. 위에서 생성한 work.quiz4 데이터셋을 이용하여 Logistic regression 분석을 수행하고 다음의 Table 4를 완성하십시오.
(회색 글씨는 표시 형식을 보여주는 예이므로 삭제하고, 실제 숫자를 확인후 검은색 글씨로 채워 넣으시오)

- 종속변수(dependent variable): hypothyroidism (이분형 변수)
- 독립변수(independent variable): 혈중카드뮴농도 (연속형 변수, 단위 $\mu\text{g/L}$)
- 교란변수(confounding variable): age(연속형 변수), BMI(연속형 변수), smoking status(이분형 변수, 현재흡연여부), UI/Cre(연속형 변수), TPOAb(연속형 변수)

Table 4. Association between blood cadmium and hypothyroidism

	Hypothyroid status	
	Men (n=1055)	Women (n=896)
BCd	OR (95% CI)	OR (95% CI)
Model 1	0.543 (0.363-0.811)	1.184 (0.923-1.518)
Model 2	0.677(0.373-1.232)	1.307 (0.946-1.805)
Model 3	0.665(0.363-1.218)	1.306 (0.942-1.811)

Model 1: unadjusted

Model 2: adjusted for age, BMI, smoking status

Model 3: Model 2 plus additionally adjusted for UI/Cre, TPOAb

Bcd, blood cadmium; OR, Odds ratio; CI, Confidence Interval; BMI, body mass index; UI/Cre, urine iodine-to-creatinine ratio; TPOAb, anti-thyroid peroxidase antibody.

[SAS code]

```
PROC SORT DATA=work.quiz4;
BY sex;
RUN;
/*Model 1*/
proc logistic data=work.quiz4;
    model hypothyroidism(event='1') = HE_Cd;
    by sex;
run;

/*Model 2*/
proc logistic data=work.quiz4;
    model hypothyroidism(event='1') = HE_Cd age HE_BMI sm_presnt;
    by sex;
run;

/*Model 3*/
proc logistic data=work.quiz4;
    model hypothyroidism(event='1') = HE_Cd age HE_BMI sm_presnt UICre HE_TPOAb;
    by sex;
run;
```