

## 7장. 자료의 정리2 (이변량 데이터)

### 7.3 범주형자료의 이원분할표

#### (1) 예제자료

안전벨트 착용 유무	아이 착용	아이 미착용
부모착용	54	7
부모미착용	3	12

```

x = rbind(c(54,7) , c(3,12) )           #자료입력
rownames(x) = c("부모착용", "부모미착용") #행이름 입력
colnames(x) = c("아이착용", "아이미착용")  #열이름 입력

margin.table(x,1)                        # 행별 도수
margin.table(x,2)                        # 열별 도수

addmargins(x)                            # 이원분할표(도수)
prop.table(x)                            # 이원분할표(비율)

barplot(x, main = "벨트착용유무", legend.text=T)           #막대그래프
barplot(x, main = "벨트착용유무", legend.text=T, beside = T) #막대그래프

```

#### (2) 우리자료

a) 그냥해도 되면 아래 b) 생략

b) [파일] => [작업공간 불러오기] # data1.RDATA => 불러오기

```

x=table(gender, eco)           # 행렬로 만들기...순서주의...=> 수정

addmargins(x)                  # 이원분할표(도수)
prop.table(x)                  # 이원분할표(비율)

barplot(x, main = "성별 경제력", legend.text=T)           #막대그래프
barplot(x, main = "성별 경제력", legend.text=T, beside = T) #막대그래프

```

## 7.4 상관계수( 둘다 양적자료 )

### (1) 예제자료

```
machine = c(68, 82, 94, 106, 92, 80, 76, 74, 110, 93, 86, 65, 74, 84, 100)
```

```
expert = c(72, 84, 89, 100, 97, 88, 84, 70, 103, 84, 86, 63, 69, 87, 93)
```

```
plot(machine, expert)
```

```
# 산점도
```

```
cor(machine, expert)
```

```
# Pearson's 상관계수
```

```
cor(machine, expert, method = "spearman")
```

```
# Spearman's 상관계수
```

```
cor.test(machine, expert)
```

```
# 가설검정  $H_0 : \rho = 0$ 
```

### 95% Pearson's sample correlation coefficient 프로그래밍

- Pearson's 표본상관계수(Pearson's sample correlation coefficient)

$$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{\frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}}{\sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}}} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}$$

#### a) 쉽게

```
r1 = cov(machine, expert) / ( sd(machine) * sd(expert)) ; r1 # 프로그래밍
```

#### b) 반복문 사용

```
x= machine ; y = expert
```

```
xbar = mean(x) ; ybar = mean(y)
```

```
sum1 = 0 ; sum2 = 0 ; sum3 = 0
```

```
for ( i in 1 : length(x)) {
```

```
    sum1 = sum1 + (x[i] - xbar)^2
```

```
    sum2 = sum2 + (y[i] - ybar)^2
```

```
    sum3 = sum3 + (x[i] - xbar)*(y[i] - ybar)
```

```
}
```

```
r2 = sum3 / sqrt( sum1* sum2) ; r2
```

## (2) 우리자료

```
plot(height, weight)          # 산점도
cor(height, weight)           # Pearson's 상관계수
r = cov(height, weight) / ( sd(height) * sd(weight)) ; r # 프로그래밍

cor(height, weight, method = "spearman")    # Spearman's 상관계수
cor.test(height, weight)                    # 가설검정  $H_0 : \rho = 0$ 

plot(mid, final)
cor(mid, final)
cor(mid, final, method = "spearman")
cor.test(mid, final)
```

**[과제20]** 일단 따라서 한번 해보시고, 아래 자료로 실습하고 해석도 해보시길..

### 1) 범주형자료의 이원분할표

	주름있음	주름없음
흡연	190	110
비흡연	206	464

### 2) 성별, 주관적 경제력

### 3) IQ와 평균성적으로 연습하시길

첨부파일 : 학번이름20.hwp (예 : 20192260홍길동20.hwp)