

10. 통계적 분석(t-test vs. ANOVA) # src: 7_R/10_t-test_ANOVA분석 참고1) **t-test분석**; 두 그룹간 평균의 차이가 통계적으로 유의한지 검증

- (1) 도표로 차이의 유의성 추출(`tapply`, `by`, `summaryBy`, `aggregate`, ... 평균, 분산 등)
- (2) 그래프 등 시각화(바이올린 도표(`geom_violin`), 산점도, 상자 도표(`geom_boxplot`) 등)

(3) 통계적 분석; **p-value**가 0.05보다 작은 경우 ⇒ 귀무가설 기각

- ① 일원표본 t-test; `t.test(a그룹, b그룹)`, `t.test(a그룹-b그룹)`
- ② 이원표본 t-test; ①등분산성(분산이 같은지) 테스트 → ②t.test 분석

```
var.test(종속변수~독립변수, data=자료명)
```

#p-value가 0.05보다 작은 경우 귀무가설(등분산성) 기각 ⇒ 등분산성 성립X (`var.equal=F`)

#p-value가 0.05보다 큰 경우 귀무가설 기각할 수 없음 ⇒ 등분산성 성립O (`var.equal=T`)

```
t.test(종속변수~독립변수, data=자료명, var.equal=T or F)
```

#p-value가 0.05보다 작은 경우: 귀무가설(두 그룹의 평균이 같다)을 기각

#p-value가 0.05보다 큰 경우: 귀무가설을 기각할 수 없음

2) **ANOVA분석(분산분석)**; 세 그룹 이상 집단의 평균 차이가 통계적으로 유의한지 검증

- (1) 도표로 평균 차이 추출
- (2) 평균, 표본의 차이 시각화

(3) 통계적 분석; **F값**이 0.05보다 작은 경우 ⇒ 귀무가설 기각

- ① `aov()`; 독립변수가 범주형일 때 주로 사용

```
result <- aov(종속변수~독립변수, data=자료명)
```

`summary(result)` #F값<0.05: 귀무가설 기각, #F값>0.05: 귀무가설 기각할 수 없음

- ② `anova()`; 독립변수가 연속형일 때 주로 사용

```
fit <- lm(종속변수~독립변수, data=자료명)
```

`anova(fit)` #F값<0.05: 귀무가설 기각, #F값>0.05: 귀무가설 기각할 수 없음