Chapter 2. Analysis of Variance (ANOVA)

2.1 Two-Sample t-Tests in PROC TTEST

Objectives

- Analyze differences between two population means using the TTEST procedure.
- Verify the assumptions of a two-sample t-test.
- Perform a one-sided test.

➔ two sample t-test는 두 그룹 간에 평균 차이가 있다고 할 수 있는 지에 대한 가설검정을 하기 위해 사용한다.

Assumptions

- independent observations
- normally distributed data for each group
- equal variances for each group

F-test for Equality of Variances(등분산성 검정)

To evaluate the assumption of equal variances in each group you can use graphics or the F-test for equality of variances. The *null hypothesis* for this test is that the *variances are equal*. When performing this test, note that if the null hypothesis is true, *F* tends to be close to 1.

$$\mathrm{H0} : {\sigma_1}^2 = {\sigma_2}^2$$
$$\mathrm{H1} : {\sigma_1}^2 \neq {\sigma_2}^2$$

$$\mathrm{F} = \frac{max({s_1}^2, {s_2}^2)}{min({s_1}^2, {s_2}^2)}$$

➔ F값이 커질수록 귀무가설을 기각할 확률이 높아진다.

The TTEST Procedure

```
PROC TTEST DATA=SAS-data-set <options>;
      CLASS variable;
      VAR variables;
RUN;
```

CLASS specifies the two-level variable for the analysis. Only one variable is allowed in the **CLASS** statement.

VAR specifies numeric response variables for the analysis. If the **VAR** statement is not specified, **PROC TTEST** analysis all numeric variables in the input data set that are not listed in a **CLASS** (or **BY**) statement.

Equal Variance t-Test and p-Values

① t-Tests for Equal Means: H0: $\mu_1 - \mu_2 = 0$

Equal Variance t-Test(Pooled):

T = 7.4017 DF = 6.0 Prob > $|T|$ = 0.0003

Unequal Variance t-Test(Satterthwaite):

T = 7.4017 DF = 5.8 Prob > $|T|$ = 0.0004

② F-test for Equal Variance: H0: $\sigma_{12} = \sigma_{22}$

Equality of Variances Test(Folded F):

F' = 1.51 DF = (3,3) Prob > F' = 0.7446

➔ ② First, check the assumption for equal variances and then use the appropriate test for equal means. Because the p-value of the test F-statistic if 0.7446, there is not enough evidence to reject the null hypothesis of equal variances.

귀무가설 기각 안함. 등분산성을 만족한다.

➔ Therefore, ① use the equal variance t-test line in the output to test whether the means of the two populations are equal.

귀무가설 기각. 두 집단의 모평균은 같지 않다.

Unequal Variance t-Test and p-Values

① t-Tests for Equal Means: H0: $\mu_1 - \mu_2 = 0$

Equal Variance t-Test(Pooled):

T = -1.7835 DF = 13.0 Prob > $|T|$ = 0.0979

Unequal Variance t-Test(Satterthwaite):

T = -2.4518 DF = 11.1 Prob > $|T|$ = 0.0320

② F-test for Equal Variance: H0: $\sigma_{12} = \sigma_{22}$

Equality of Variances Test(Folded F):

F' = 15.28 DF = (9,4) Prob > F' = 0.0185

➔ ② Again, first check the assumption for equal variances and use the approximate test for equal means. Because the p-value of the F-statistic is less than alpha=0.05, there is enough evidence to reject the null hypothesis of equal variances.

귀무가설 기각. 등분산성을 만족하지 않는다.

➔ Therefore, ① use the unequal variance t-test line in the output to test whether the means of the two populations are equal.

귀무가설 기각. 두 집단의 모평균은 같지 않다.

One Sided Tests and Confidence Intervals

In many situations one might decide that rejection on only one side of the mean is important. For instance, a drug company might only want to test for positive differences between their new drug and placebo and not negative differences. One-sided tests are a way of going about doing this.

- Used when the null hypothesis is of the form:

  H0: $\mu \le k$ or H0: $\mu \ge k$

- Can increase power.

- Tests and Confidence Intervals produced in PROC TTEST using:

  - SIDES=U for Upper Tail Tests ($\mu \le k$) and CIs

  - SIDES=L for Lower Tail Tests ($\mu \ge k$) and Cis

2.2 One-Way ANOVA(일원 분산분석)

Overview

*Analysis of variance(*ANOVA) is a statistical technique used to compare the means of two or more groups of observations or treatments. For this type of problem, you have a

- continuous dependent variable, or response variable(반응변수, 종속변수)
- discrete independent variable also called a predictor or explanatory variable(설명변수, 독립변수)

➔ 3개 이상의 그룹의 모평균을 비교하는 경우

Research Questions for One-Way ANOVA

- Do accounts, on average, earn more than teachers?
  - ➔ a t-test can be thought of as a special case of ANOVA: if you analyze the difference between means using ANOVA, you get the same results as with a t-test.
- Do people treated with one of two new drugs have higher average T-cell counts than people in the control group?
- Do people spend different amounts depending on which type of credit card they have?
- Does the type of fertilizer used affect the average weight of garlic grown at the Montana Gourmet Garlic ranch?

Assumptions for ANOVA

- Observations are independent.
- Errors are normally distributed.
- All groups have equal response variances.

Garlic Example

Montana Gourmet Garlic, a company that grows garlic using organic method design an experiment to test whether growth of the garlic is affected by the type of fertilizer used. They test three different organic fertilizers and one chemical fertilizer (as a control).

The variables in the data set are

Fertilizer	The type of fertilizer used (1 through 4).
BulbWt	The average garlic bulb weight (in pounds) in the bed.
Cloves	The average number of cloves on each bulb.

BedID A randomly assigned bed identification number.

The ANOVA Model

$$Y_{ik} = \mu + \tau_i + \epsilon_{ik}$$

BulbWt = Base Level + Fertilizer + Unaccounted for Variation

Y_{ik} the k^{th} value of the response variable for the i^{th} treatment.

μ the overall population mean of the response, for instance garlic bulb weight

τ_i the difference between the population mean of the i^{th} treatment and the overall mean μ. This is referred to as the effect of treatment i.

ϵ_{ik} the difference between the observed value of the k^{th} observation in the i^{th} group. This is called the error term.

Partitioning Variability in ANOVA

Model Sum of Squares (SSM)

: the variability explained by the independent variable and therefore represented by the between treatment sums of squares.

Error Sum of Squares (SSE)

: the variability not explained by the independent variable. Also referred to as within treatment variability or residual sum of squares.

Total Sum of Squares (SST)

: the overall variability in the response variable.

SST = SSM + SSE

➔ SSM 이 클수록 그룹간 차이가 큼을 의미한다.

The GLM Procedure

PROC GLM DATA=*SAS-data-set* PLOTS=*options*;
    CLASS *variables*;
    MODEL *dependents=independents* </ *options*>;
    MEANS *effects* </ *options*>;
    LSMEANS *effects* </ *options*>;
    OUTPUT OUT=*SAS-data-set keyword=variable...*;
RUN;
QUIT;

CLASS specifies classification variables for the analysis

MODEL specifies dependent and independent variables for the analysis

MEANS computes unadjusted means of the dependent variable for each value of the specified effect

LSMEANS produces adjusted means for the outcome variable, broken out by the variable specified and adjusting for any other explanatory variables included in the **MODEL** statement.

OUTPUT specifies an output data set that contains all variables from the input data set and variables that represent statistics from the analysis

.Assessing ANOVA Assumptions

- Good data collection methods help ensure the independence assumption
- Diagnostic plots from PROC GLM can be used to verify the assumption that error is approximately normally distributed.
- The GLM procedure produces a hypothesis test with the HOVTEST option in the MEANS statement. H0 for this hypothesis test is that the variances are equal for all populations.

2.3 ANOVA with Data from a Randomized Block Design

Observational or Retrospective Studies (관찰연구 또는 후향적 연구)

- Groups can be naturally occurring.

  - for example, gender and ethnicity.

- Random assignment might be unethical or untenable.

  - for example, smoking or credit risk groups.

- Often you look at what already happened (retrospective) instead for following through the future (prospective).

- You have little control over other factors contributing to the outcome measure.

- 공중보건이나 경영분석 연구

Controlled Experiments (전향적 연구)

- Random assignment might be desirable to eliminate selection bias.

- You often want to look at the outcome measure prospectively.

- You can manipulate the factors of interest and can more reasonably claim causation.

- You can design your experiment to control for other factors contributing to the outcome measure.

- 설계 시 요인 통제 가능

Assigning Treatments within Blocks

마늘 예제를 보면 일조량, 토양의 산성도 등의 잡음 요인이 존재하는데 이들은 통제가 불가능하다. 잡음을 줄이기 위해 sector를 나눠서 한 sector 내에서 비료를 무작위로 배정한다.

randomized block design (난괴설계: 무선화 구획설계)

: 종속변수와 관련이 높은 구획변수를 통제함으로써 독립변수의 종속변수에 대한 영향력을 정확히 측정하기 위해 사용하는 기법

$$Y_{ijk} = \mu + \alpha_i + \tau_j + \epsilon_{ijk}$$

BulbWt = Base Level + Sector + Fertilizer + Unaccounted for Variation

2.4 ANOVA Post Hoc Test (분산분석의 사후 검정; 다중 비교)

Objectives

- Perform pairwise comparisons among groups after finding a significant effect of an independent variable in ANOVA.
- Demonstrate graphical features in PROC GLM for performing post hoc test.
- Interpret a diffogram.
- Interpret a control plot.

Multiple Comparison Methods

Comparison-wise Error Rate (α = 0.05)	Number of Comparisons	Experiment-wise Error Rate (α = 0.05)
0.05	1	0.05
0.05	3	0.14
0.05	6	0.26
0.05	10	0.40

When you control the comparison-wise error rate (CER), you fix the level of alpha for a single comparison, without taking into consideration all the pairwise comparisons you are making.

The experiment-wise error rate (EER) uses an alpha that takes into consideration all the pairwise comparisons you are making. Presuming no differences exist, the chance you falsely conclude at least one difference exists is much higher when you consider all possible comparisons.

If you want to make sure that the error rate is 0.05 for the entire set of comparisons, use a method that controls the experiment-wise error rate at 0.05.

<table>
<tr><td>Control Comparison-wise Error Rate</td><td>Pairwise t-tests</td></tr>
<tr><td rowspan="2">Control Experiment-wise Error Rate</td><td>Compare All Pairs
Tukey</td></tr>
<tr><td>Compare to Control
Dunnett</td></tr>
</table>

- Control Comparison-wise Error Rate
  - ✓ fix the level of alpha for a single comparison, without taking into consideration the number of pairwise comparisons you are making.
  - ✓ LSMEANS / PDIFF=ALL ADJUST=T;
- Control Experiment-wise Error Rate

- ✓ use an alpha that takes into consideration all the pairwise comparisons you are making.
- ✓ LSMEANS / PDIFF=ALL ADJUST=TUKEY;
- ✓ LSMEANS / PDIFF=CONTROL() ADJUST=DUNNETT;

<u>Tukey's Multiple Comparison Method</u>

This method is appropriate when considering pairwise comparisons only.

The experiment-wise error rate is

- equal to alpha when all pairwise comparisons are considered
- less than alpha when fewer that all possible comparisons are considered
  - ✓ If you choose to do fewer than all pairwise comparisons, then this method is more conservative.
- Diffogram - used to quickly tell if two group means are statistically significant.

Dunnett's Multiple Comparison Method

This method is appropriate when comparing to a control group, such as a placebo group in a drug trial.

- Experiment-wise error rate is no greater than the stated alpha.
- Takes into account the correlations among trials.
- One-sided hypothesis tests against a control group can be performed.
- Computes and tests k-1 groupwise differences, where k is the number of levels of the CLASS variable.
- Recommended when there is a true control group.
- Control plots

2.5 Two-Way ANOVA with Interactions (이원 분산분석)

n-Way ANOVA

- A case with two categorical predictors.
- In general, any time you have more than one categorical predictor variable and a continuous response variable, it is called *n*-way ANOVA. The *n* can be replaced with the number of categorical predictor variables.
- Randomized block design is actually a special type of *n*-way ANOVA.

Drug Example

The purpose of the study is to look at the effect of a new prescription drug on blood pressure. Data was collected in an effort to determine whether different dose levels of a given drug have an effect on blood pressure for people with one of three types of heart disease.

- Response variable: Blood pressure
- Explanatory variable: Drug Dose (100mg, 500mg, Placebo)

  Disease (A, B, C)

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

BloodP = Base Level + Disease + Drug Dose + DrugDose and Disease + Unaccounted for Variation

Interactions

When you analyze an *n*-way ANOVA with interactions you should first look at any tests for interaction among factors.

If there is no interaction between the factors, the test for the individual factor effects can be interpreted to determine their significance/non-significance.

If an interaction exists between any factors, the tests for the individual factor effects might be misleading, due to masking of these effects by the interaction. This is especially true for unbalanced data.

➔ 교호작용이 유의하지 않은 경우, 아예 교호작용 항을 제거하고 분석하는 방법도 있다.