Lecture 21

Part 4 Analysis of Variance

Relationship between Chi-squared Test and F Test

Another Way to Define Chi-squared Distribution and Test

- Let $Z_1, Z_2, \ldots, Z_k$ being independent and identically distributed and follow $N(0,1)$

$$\Rightarrow X^2 \equiv Z_1^2 + Z_2^2 + \ldots + Z_k^2 \sim \chi_k^2.$$

- Specifically, if $k = 1$,

$$Z^2 \sim \chi_1^2.$$

- Chi-squared statistic: $\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$
- Where:
  - O_i are observed frequencies
  - E_i are expected frequencies
- Under H_0: $\chi^2 \sim \chi_{k-1}^2$

Recall How F-test is Defined

- F-test statistic:

$$F = \frac{S_1^2/(k_1 - 1)}{S_2^2/(k_2 - 1)}$$

- Where:
  - S_1^2, S_2^2 are sample variances
  - $k_1 - 1, k_2 - 1$ are degrees of freedom
- Under H_0: $F \sim F_{k_1-1,k_2-1}$

Mathematical Connection

- Key relationship here:

$$F = \frac{\chi_1^2/(k_1 - 1)}{\chi_2^2/(k_2 - 1)}$$

- Where:
  - $\chi_1^2 \sim \chi_{k_1-1}^2$
  - $\chi_2^2 \sim \chi_{k_2-1}^2$
  - χ_1^2 and χ_2^2 are independent

Fisher's F-statistic

F-statistics is invented by and named after Sir Ronald Fisher.

The original concern of Fisher is to construct a **statistic** which has a sampling distribution, in some extent, free from the degrees of freedom a and b *under the null hypothesis*.

Fisher's F-statistic

With this concern, he presented his F-statistic in a way that:

Since χ_a^2 has expectation a, so the numerator χ_a^2/a has expectation 1;

similarly, the denominator also has expectation 1.

As Fisher said, **the value of F-statistic will fluctuate near 1 under the null hypothesis** $H_0 : \mu_1 = \cdots = \mu_k$ (if $k = a + 1$).

Recall One-Way ANOVA We Have Learned

In one-way ANOVA:

$$F = \frac{\text{Between-group variability}}{\text{Within-group variability}} = \frac{\chi^2_{\text{between}}/(k-1)}{\chi^2_{\text{within}}/(N-k)}$$

Where:

- k is number of groups
- N is total sample size

Applications

- Variance Comparison
  - F-test for comparing two variances
  - Multiple χ^2 tests for multiple variances
- Model Comparison
  - F-test in regression analysis
  - χ^2 test for nested models

Practical Implications

- Understanding the relationship helps in:
  - Test selection
  - Result interpretation
  - Statistical power considerations

Relationship between t Test and F Test

Relationship between t Test and F Test

- All under the assumption that null hypothesis is true

$$
\begin{aligned}
t_{k-1} &= \frac{Z}{\sqrt{S^2/(k-1)}} \\
&= \frac{Z}{\sqrt{\chi^2_{k-1}/(k-1)}} \\
&= \frac{\sqrt{\chi^2_1/1}}{\sqrt{\chi^2_{k-1}/(k-1)}} = \sqrt{\frac{\chi^2_1/1}{\chi^2_{k-1}/(k-1)}} \\
&= \sqrt{F_{1,k-1}}
\end{aligned}
$$

Or, in other words,

$$t^2_{k-1} = F_{1,k-1}.$$