

**Class Notes: STAT 501**

**Nonparametrics & Log-Linear Models**

**Ansari-Bradley Test, Kruskal-Wallis Test**

Da Kuang  
University of Pennsylvania

# Contents

<b>1</b>	<b>Ansari-Bradley Test</b>	<b>4</b>
1.1	Background . . . . .	4
1.1.1	$\chi^2$ Distribution and $F$ Distribution . . . . .	4
1.1.2	$F$ -test of Equality of Variances . . . . .	4
1.2	Assumption . . . . .	5
1.3	Hypothesis Test . . . . .	6
1.3.1	Assign Scores . . . . .	7
1.3.2	Two-sample Scale Statistic . . . . .	7
<b>2</b>	<b>Kruskal-Wallis Test</b>	<b>9</b>
2.1	Background . . . . .	9
2.1.1	The One-Way Layout . . . . .	9
2.1.2	One-Way ANOVA . . . . .	10
2.2	Assumption . . . . .	11
2.3	Hypothesis . . . . .	12
2.3.1	Rank . . . . .	12
2.3.2	Test Statistic . . . . .	12
2.4	Extensions . . . . .	14

2.4.1	Jonckheere-Terpstra Test . . . . .	14
2.4.2	Mack-Wolfe Test . . . . .	14

# 1 Ansari-Bradley Test

## 1.1 Background

### 1.1.1 $\chi^2$ Distribution and $F$ Distribution

The  $\chi^2$  distribution is obtained directly from independent, standard normal random variables. Let  $Z_i, i = 1, 2, \dots, n$ , be  $n$  independent random variables, each distribution as standard normal. Define a new random variable as the sum of the squares of the  $Z_i$ :

$$X = \sum_{i=1}^n Z_i^2$$

Then,  $X$  has what is known as a  $\chi^2$  distribution with  $n$  degree of freedom.

Another important distribution for statistics and econometrics is the  $F$  distribution. In particular, the  $F$  distribution will be used for testing hypotheses in the context of multiple regression analysis.

To define an  $F$  random variable, Let  $X_1 \sim \chi_{k_1}^2$  and  $X_2 \sim \chi_{k_2}^2$  and assume that  $X_1$  and  $X_2$  are independent. Then the random variable

$$F = \frac{X_1/k_1}{X_2/k_2}$$

has an  $F$  distribution with  $(k_1, k_2)$  degree of freedom. We denote this as  $F \sim F_{k_1, k_2}$ .

### 1.1.2 $F$ -test of Equality of Variances

In statistics, an  $F$ -test of equality of variances is a test for the null hypothesis that two normal populations have the same variance.

- Let  $X_1, \dots, X_n$  and  $Y_1, \dots, Y_m$  be iid samples from two populations which each has a normal distribution.
- The expected values for the two populations can be different.
- The Hypothesis is that the variances are equal.

Sample means are

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \bar{Y} = \frac{1}{m} \sum_{i=1}^m Y_i.$$

Sample variances are

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2, S_Y^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i - \bar{Y})^2.$$

Then the test statistic is an  $F$ -distribution with  $n-1$  and  $m-1$  degrees of freedom if the null hypothesis of equality of variances is true.

$$F = \frac{S_X^2}{S_Y^2}$$

Otherwise it follows an  $F$ -distribution scaled by the ratio of true variances. The null hypothesis is rejected if  $F$  is either too large or too small based on the desired alpha level

## 1.2 Assumption

Ansari-Bradley Test is a non-parametric version of  $F$ -test. The null hypothesis of interest here is that the  $X$  and  $Y$  variable have the same probability distribution but their common distribution is not specified.

We have the following assumptions for Ansari-Bradley Test:

- $X_i, i \leq m$ , iid from continuous population 1 with distribution  $F$ ;
- $Y_j, j \leq n$ , iid from continuous population 2 with distribution  $G$ ;

- $X'_i$ s and  $Y'_j$ s mutually independent;
- $F$  and  $G$  have the same median  $\theta$ .

Note that  $X_i$  and  $Y_i$  are continuous random variables. Also, Ansari-Bradley test works with different medians between samples but here we assume two samples have the same median,  $\theta$ , for simplicity.

The alternative hypothesis can be that  $Y$  population has the same general form as the  $X$  population, but they could have different scales. It can be written as

$$\frac{X - \theta}{\eta_x} \stackrel{d}{=} \frac{Y - \theta}{\eta_y}$$

, where  $\stackrel{d}{=}$  means “has the same distribution”,  $\eta_x$  and  $\eta_y$  are the scale parameters of  $F$  and  $G$ , respectively.

The parameter of interest is the ratio of the scale parameters,  $\gamma = \frac{\eta_x}{\eta_y}$ .

In terms of this location-scale parameter model with equal location parameters, the null hypothesis is

$$H_0 : \gamma^2 = 1$$

, corresponding to the assertion that the population scale parameters are equal.

why  $\gamma^2$  instead of  $\gamma$ . Sometimes we cannot calculate the variance (second moment), e.g. Poisson Distribution, during standardization, so we keep the squaring later after the standardization.

## 1.3 Hypothesis Test

- $H_0 : \gamma^2 = 1$
- $H_1 : \gamma^2 > 1$  (right-tailed test)
- $H_1 : \gamma^2 < 1$  (left-tailed test)

- $H_1 : \gamma^2 \neq 1$  (two-sided test)

### 1.3.1 Assign Scores

Order the combined samples of  $N = m + n$   $X_i$ 's and  $Y_j$ 's from least to greatest.

- Assign the score 1 to both the smallest and the largest observations in this combined sample.
- Assign the score 2 to the 2nd smallest and 2nd largest.
- ...

In the end, we will get a list of score such that

- If  $N$  is an even integer, the array of assigned scores is

$$1, 2, \dots, \frac{N}{2}, \frac{N}{2}, \dots, 2, 1$$

- If  $N$  is an odd integer, the array of assigned score is

$$1, 2, \dots, \frac{N-1}{2}, \frac{N+1}{2}, \frac{N-1}{2}, \dots, 2, 1$$

### 1.3.2 Two-sample Scale Statistic

Let  $R_j$  be the score assigned to  $Y_j$ ,  $j = 1, \dots, n$ . The test statistic is  $C = \sum_{j=1}^n R_j$ .

Under the null hypothesis, we have

$$\mathbf{E}C = \frac{n(N+2)}{4}, \mathbf{Var}C = \frac{mn(N+2)}{48(N-1)}$$

The big idea is that we calculate the critical value  $c_\alpha$  based on the Type I error of the experiment design and decide whether to reject the null hypothesis by comparing  $C$  and  $c_\alpha$ . Intuitively, if  $\gamma^2 > 1$ , the  $X$  values would tend to be more spread out than the  $Y$

values. Thus, the  $Y'_j$ s would tend to get larger scores than the  $X'_i$ s. **Then  $C$  would tend to be larger.**

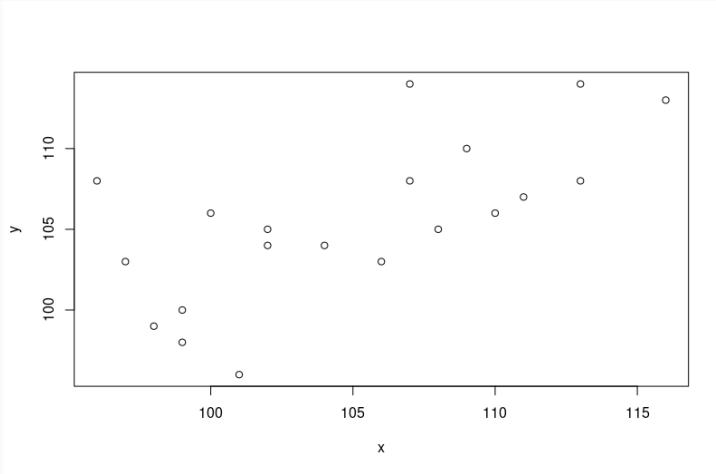
Note that  $\gamma$  shows in the hypothesis to describe the data but the test statistics is actually  $C$ , the sum of ranks in  $Y$ .

The p-value of the hypothesis test can be calculated from the build-in R function `ansari.test`.

### Listing 1: Ansari Example

```
1 x= c(111, 107, 100, 99, 102, 106, 109, 108, 104, 99,
2 101, 96, 97, 102, 107, 113, 116, 113, 110, 98)
3 y=c(107, 108, 106, 98, 105, 103, 110, 105, 104,
4 100, 96, 108, 103, 104, 114, 114, 113, 108, 106, 99)
5 plot(x,y)
6 ansari.test(x,y)
7
8 # Ansari-Bradley test
9 #
10 # data: x and y
11 # AB = 185.5, p-value = 0.1815
12 # alternative hypothesis: true ratio of scales is not equal to 1
```





**Figure 1**

Note that the range of X and Y are very similar, which indicates that they share the same dispersion. Also, in the Ansari-Bradley test, we get p-val = 0.1815, so that we cannot reject the null hypothesis.

## 2 Kruskal-Wallis Test

### 2.1 Background

#### 2.1.1 The One-Way Layout

One-way, also known as single-factor, means that there is only one treatment among the groups in data.

Denote the data consist of  $N = \sum_{j=1}^I n_i$  observations, with  $n_i$  observations from the  $i$ th treatment,  $i = 1, \dots, I$ .

1	2	...	I
$Y_{1,1}$	$Y_{2,1}$	$\cdots$	$Y_{I,1}$
$Y_{1,2}$	$Y_{2,2}$	$\cdots$	$Y_{I,2}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$
$Y_{1,n_1}$	$\vdots$	$\cdots$	$\vdots$
	$Y_{2,n_2}$	$\cdots$	$\vdots$
			$Y_{I,n_I}$

**Figure 2:** Data

Note that each group can have different number of observations.

### 2.1.2 One-Way ANOVA

In statistics, one-way analysis of variance is a technique that can be used to compare means(location) of two **or more samples**. The primary interest consist in comparing the location of three of more samples.

- Null Hypothesis: the sample data are supposed to come from the same population.
- Alternative Hypothesis: the sample data are supposed to come from different populations.

Recall in ANOVA, the variability of data is broken down into two components:

- Sum of squares of the columns: between groups variability, due to the different locations of the populations.
- Sum of squares of the error: within groups variability, due to the variability in the population.

DEGREES OF FREEDOM FOR ONE-WAY ANOVA			
$SSC$	Sum of squares (columns/treatments)	$df_{columns} = 3 - 1 = 2$	$MSC = \frac{SSC}{2}$
$SSE$	Sum of squares (within/error)	$df_{error} = 21 - 3 = 18$	$MSE = \frac{SSE}{18}$
$SST$	Sum of squares (total)	$df_{total} = 21 - 1 = 20$	$F = \frac{MSC}{MSE}$
$MSC$ = mean square columns/treatments		$MSE$ = mean square error / within	

**Figure 3:** Analysis of Variance

$F$ -ratio is the test statistic for the hypothesis test.

## 2.2 Assumption

- The random variables are all mutually independent.
- Random variables in the same group follow the same continuous distribution  $F_i$ ,  $i = 1, \dots, I$ .
- $F_i(t) = F(t - \tau_i)$ ,  $-\infty < t < \infty$ .
  - $F$  is a continuous distribution function with unknown median  $\theta$ ;
  - $\tau_i$  is the unknown treatment effect for the  $i$ -th population.
  - No difference in scale parameters (variability).

We can set up the model of each observation based on the assumptions.

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

- $i = 1, \dots, I, k = 1, \dots, n_i$ .
- $\theta$ : overall median.
- $\tau_i$ : the  $i$ -th treatment effect.

- $\epsilon_{ik}$ : random error. iid from a continuous distribution with median 0.

## 2.3 Hypothesis

- $H_0$ :  $\tau_1 = \dots = \tau_I$ , which means that  $F_1 = \dots = F_I = F$ .
- $H_1$ :  $\tau_1 \dots \tau_I$  are not all equal.

### 2.3.1 Rank

Combine all  $N$  observations from the  $I$  groups and order them from least to greatest.

Example:

- $I = 3$  groups;
- $N = \sum_{i=1}^3 n_i = 1 + 1 + 2 = 4$  observations.

	group 1	group 2	group 3
Observation	5	6	{9,10}
Rank	1	2	{3,4}

**Figure 4:** Example: Rank

### 2.3.2 Test Statistic

Let  $r_{ik}$  denotes the rank of  $Y_{ik}$  in this joint ranking. The average rank assigned in the joint ranking is

$$\frac{1}{N} \sum_{i=1}^I \sum_{k=1}^{n_i} r_{ik} = \frac{N(N+1)}{2N} = \frac{N+1}{2}$$

Let  $R_i = \sum_{k=1}^{n_i} r_{ik}$ , which is the sum of the joint ranks received by the observations in the  $i$ -th group.

Let  $\bar{R}_i = R_i/n_i, i = 1, \dots, I$ , which is the average rank received by observations in the  $i$ -th group.

The Kruskal-Wallis statistic  $H$  is then given by

$$\begin{aligned} H &= \frac{12}{N(N+1)} \sum_{i=1}^I n_i \left( \bar{R}_i - \frac{N+1}{2} \right)^2 \\ &= \left( \frac{12}{N(N+1)} \sum_{i=1}^I \frac{R_i^2}{n_i} \right) - 3(N+1). \end{aligned}$$

Reject  $H_0$  if  $H \geq h_\alpha$ , the upper  $\alpha$  percentile for the null distribution of  $H$ .

Under the Null Hypothesis,

- $\mathbf{E}r_{ik} = \frac{N+1}{2}$
- $\mathbf{E}\bar{R}_i = \frac{1}{n_i} \sum_{k=1}^{n_i} r_{ik} = \frac{1}{n_i} n_i \mathbf{E}r_{ik} = \frac{N+1}{2}$ .

Intuitively, we expect the  $\bar{R}_i$ 's to be close to  $\frac{N+1}{2}$  when  $H_0$  is true. So that small values of  $H$  represent agreement with  $H_0$ .

Otherwise, when the  $\tau$ 's are not all equal, we would expect a portion of the associated treatment average ranks to differ from their common average,  $(N+1)/2$ , with some tending to be larger and some smaller. Hence, the net result  $n_i(\bar{R}_i - \frac{N+1}{2})^2$  would lead to a large value of  $H$ . This suggests rejecting  $H_0$  in favor of  $H_1$  for large values of  $H$ .

$$\frac{N+1}{2} = 5/2.$$

group 1	group 2	group 3
1	2	{3,4}
$R_1 = 1$	$R_2 = 2$	$R_3 = 7; \frac{R_3^2}{2} = \frac{49}{2};$
$(1 - \frac{5}{2})^2 = \frac{9}{4}$	$(2 - \frac{5}{2})^2 = \frac{1}{4}$	$\bar{R}_3 = \frac{7}{2}; 2(\frac{7}{2} - \frac{5}{2})^2 = 2$

**Figure 5:** Example: Kruskal-Wallis Test

## 2.4 Extensions

Recall that the alternative hypothesis of Kruskal-Wallis Test is that  $\tau_i$ 's are not equal. Now we introduce two more test for stronger alternative hypothesis.

### 2.4.1 Jonckheere-Terpstra Test

In many practical settings, the treatments are such that the appropriate alternatives to no differences in treatment effects ( $H_0$ ) are those of increasing (or decreasing) treatment effects according to some natural labeling for the treatments.

Examples of such settings include “treatments” corresponding to degrees of knowledge of performance, quality or quantity of materials, severity of disease, amount of practice, drug dosage levels, intensity of a stimulus, and temperature.

The Kruskal-Wallis test does not utilize any such partial prior information regarding a postulated alternative ordering.

The alternative hypothesis for Jonckheere-Terpstra Test is that

$$H_1 : \tau_1 \leq \tau_2 \leq \cdots \leq \tau_I$$

, with at least one strict inequality.

### 2.4.2 Mack-Wolfe Test

The alternative hypothesis for Mack-Wolfe Test is that

$$H_1 : \tau_1 \leq \cdots \leq \tau_p \geq \tau_{p+1} \geq \cdots \tau_I$$

, with at least one strict inequality.

The distribution of  $\tau_i$ 's are in the shape of umbrella. But this test rarely used nowadays.