Statistical Methods

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

December 6, 2019

Notes

Topics

Notes

Objectives

Objectives:

- Carry out a Kruskal-Wallis test for I population means.

Notes

Kruskal-Wallis Test for I Population Means $\mu_1, \mu_2, \ldots, \mu_I$

- The ***Kruskal-Wallis test*** is a **nonparametric** alternative to the **one-factor ANOVA F test** for comparing I population means $\mu_1, \mu_2, \ldots, \mu_I$.
- We assume that we have *independent* random samples from I continuous populations that all have the **same shape** but possibly different means $\mu_1, \mu_2, \ldots, \mu_I$.

Notes

- The **null hypothesis** is that there are no differences among the population means $\mu_1, \mu_2, \ldots, \mu_I$:

Null Hypothesis:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_I$$

Notes

- The **alternative hypothesis** is that there's *at least one difference* among the set of means:

Alternative Hypothesis: The alternative hypothesis will be

H_a : At least two of the μ_i's are different

These are the same hypotheses as the ones tested in a *one-factor ANOVA F test*.

Notes

- **Notation**:

$\boldsymbol{I}$ = The number of populations from which samples are taken.

$\boldsymbol{Y_{ij}}$ = The jth observation in the ith sample.

$\boldsymbol{J_1, J_2, \ldots J_I}$ = The sample sizes (not necessarily equal).

$\boldsymbol{N}$ = The total number of observations (in the I samples combined), i.e.

$$N = \sum_i J_i .$$

Notes

- (cont'd)

Now consider **combining** the I samples and and **ranking** the observations from smallest (rank = 1) to largest (rank = N).

$\boldsymbol{R_{ij}}$ = The rank of Y_{ij} .

$\boldsymbol{\bar{R}_i}$ = The average of the ranks of the observations from the ith sample.

$\boldsymbol{\bar{R}}$ = The overall average of all N ranks.

For **ties**, use the **average** of the **ranks** that would've been assigned if there weren't any ties.

Notes

Proposition

The overall average of all N ranks is

$$\bar{R} = \frac{N+1}{2}$$

because the sum of the ranks is

$$1 + 2 + \cdots + N = \frac{N(N+1)}{2}.$$

Notes

Kruskal-Wallis Test Statistic:

$$K = \frac{12}{N(N+1)} \sum_{i=1}^{I} J_i \left(\bar{R}_i - \bar{R}\right)^2 .$$

- K is the **sum of** (J_i-*weighted*) **squares** of the **mean ranks** away from the **overall mean rank**, multiplied by the constant $12/N(N+1)$ (explained later).

  It's analogous to the **treatment sum of squares SSTr** in **one-factor ANOVA**.

Notes

- K will be **small** when the **mean ranks** $\bar{R}_1, \bar{R}_2, \ldots, \bar{R}_I$ are approximately **equal**, i.e. if the combined, sorted samples "intermingle", as would be the case if $\mu_1, \mu_2, \ldots, \mu_I$ were equal.

  K will be **large** when the **mean ranks** $\bar{R}_1, \bar{R}_2, \ldots, \bar{R}_I$ **differ**, i.e. if the combined, sorted samples "segregate", as would be the case if there were differences among $\mu_1, \mu_2, \ldots, \mu_I$

Notes

Large values of K provide **evidence against H_0 in favor of H_a : At least two of the μ_i's are different**.

Notes

- Now suppose the sample sizes $J_1, J_2, \ldots, J_I$ are all *large*.

  In this case, the sampling distribution of the test statistic is as follows.

Sampling Distribution of the Test Statistic Under H_0:
If K is the Kruskal-Wallis test statistic, then when

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_I$$

is true,

$$K \sim \chi^2(I-1)$$

(approximately), a chi-squared distribution with $(I-1)$ degrees of freedom.

Notes

- The $\chi^2(I-1)$ curve gives us:
  - The ***rejection region*** as the **extreme largest 100α% of K values**.
  - The ***p-value*** as the **tail area to the right of the observed K value**.

Notes

- **Comment**: The degrees of freedom is $I-1$ because the I (J_i-*weighted*) deviations $J_i(\bar{R}_i - \bar{R})$ sum to zero, so only $I-1$ of them are "free to vary".
- **Comment**: The constant $12/N(N+1)$ "rescales" the sum $\sum_i J_i \left(\bar{R}_i - \bar{R}\right)^2$ just enough to force K to follow a $\chi^2(I-1)$ distribution (under H_0).

Notes

Example

An agricultural experiment was carried out to examine the effects of **four** soil **treatments** on the soil **phosphorus** levels.

Twenty plots of land were randomly assigned to receive one of the **four treatments**, with **five** plots per treatment.

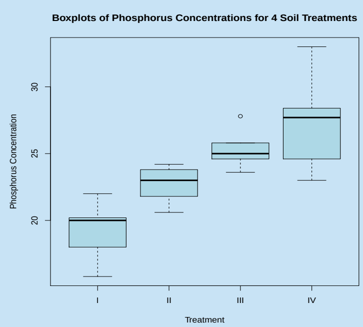
The phosphorus concentrations (mg/g) in the topsoils of the plots are shown on the next slide.

Notes

Treatment I	Treatment II	Treatment III	Treatment IV
20.2	23.0	23.6	23.0
15.8	21.8	27.8	33.0
18.0	24.2	25.8	28.4
20.0	20.6	24.6	24.6
22.0	23.8	25.0	27.7

Side-by-side **boxplots** of the data are on the next slide.

Notes



Notes

Because the four sample sizes are small, it is difficult to ascertain from plots alone whether the data are normally distributed.

Suppose we're unwilling to assume normality because previous studies have shown that soil phosphorus concentrations follow **right skewed** distributions.

Notes

We'll carry out a **Kruskal-Wallis test** of

$$H_0: \quad \mu_1 = \mu_2 = \cdots = \mu_I$$
H_a: At least two of the μ_i's are different

to decide if there are *any* significant **differences** in the mean **phosphorus** concentrations for the **four treatments**.

Notes

Here are the samples **combined**, **sorted**, and **ranked**.

Observation	15.8	18.0	20.0	20.2	20.6	21.8	22.0	23.0	23.0	23.6
Sample	I	I	I	I	II	II	I	II	IV	III
Rank	1	2	3	4	5	6	7	8.5	8.5	10

Observation	23.8	24.2	24.6	24.6	25.0	25.8	27.7	27.8	28.4	33.0
Sample	II	II	III	IV	III	III	IV	III	IV	IV
Rank	11	12	13.5	13.5	15	16	17	18	19	20

Notes

The $I = 4$ **sample sizes** are all the same:

$$J_1 = J_2 = J_3 = J_4 = 5.$$

The **overall sample size** is $N = 20$.

Notes

The **group mean *ranks*** are:

$$\bar{R}_1 = \frac{1+2+3+4+7}{5} = \mathbf{3.4}.$$

$$\bar{R}_2 = \frac{5+6+8.5+11+12}{5} = \mathbf{8.5}.$$

$$\bar{R}_3 = \frac{10+13.5+15+16+18}{5} = \mathbf{14.5}.$$

$$\bar{R}_4 = \frac{8.5+13.5+17+19+20}{5} = \mathbf{15.6}.$$

Notes

The **overall mean *rank*** is:

$$\bar{R} = \frac{20+1}{2} = \mathbf{10.5}.$$

Notes

The **test statistic** is:

$$\begin{aligned} K &= \frac{12}{N(N+1)} \sum_{i=1}^{I} J_i \left(\bar{R}_i - \bar{R}\right)^2 \\ &= \frac{12}{20(20+1)} \Big[5\left(3.4 - 10.5\right)^2 + 5\left(8.5 - 10.5\right)^2 \\ &\qquad + 5\left(14.5 - 10.5\right)^2 + 5\left(15.6 - 10.5\right)^2 \Big] \\ &= \mathbf{13.75}. \end{aligned}$$

Notes

From **Table A11** (the **chi-squared distribution** table) with $\boldsymbol{I - 1 = 3}$ **df**, the **p-value** is **between** than **0.001** and **0.005**.

Using $\alpha = 0.05$, we **reject** $\boldsymbol{H_0}$. There are differences among the phosphorus levels for the four treatments.

Notes

Notes

Notes