

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike License](https://creativecommons.org/licenses/by-nc-sa/4.0/). Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.



Copyright 2006, The Johns Hopkins University and Karl W. Broman. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided "AS IS"; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.

ANOVA, still

$\{Y_{ti}\}$ independent with $Y_{ti} \sim \text{normal}(\mu_t, \sigma)$ for $t = 1 \dots k$.

Test $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$

The usual statistic:

$$F = M_B/M_W = \frac{\sum_t n_t (\bar{Y}_t - \bar{Y}_{..})^2/k}{\sum_t \sum_i (Y_{ti} - \bar{Y}_t)^2 / (\sum_t n_t - k)}$$

P-values: (a) Use the $F(k, \sum n_t - k)$ distribution.
(b) Use a permutation test.

Assumptions: (a) Underlying dist'ns are normal with common SD.
(b) Underlying dist'ns are the same.

Non-parametric ANOVA

An alternative approach: the Kruskal-Wallis test.

Rank all of the observations from 1, 2, ..., N.

Let R_{ti} = the rank for observation Y_{ti} .

Let $\bar{R}_t = \sum_i R_{ti}/n_t$ = the average rank for group t.

Null hypothesis, H_0 : the underlying distributions are all the same.

$$E(\bar{R}_t \mid H_0) = \frac{N+1}{2}$$

$$SD(\bar{R}_t \mid H_0) = \sqrt{\frac{(N+1)(N-n_t)}{12 n_t}}$$

Kruskal-Wallis test statistic

$$H = \sum_t \left(\frac{N - n_t}{N} \right) \times \left[\frac{\bar{R}_{t.} - E(\bar{R}_{t.} | H_0)}{SD(\bar{R}_{t.} | H_0)} \right]^2$$
$$= \dots = \frac{12}{N(N+1)} \sum_t n_t \left[\bar{R}_{t.} - \left(\frac{N+1}{2} \right) \right]^2$$

Under H_0 , and if the sample sizes are large, $H \sim \chi^2(df = k - 1)$.

Alternatively, we could use a **permutation test** to estimate a P-value.

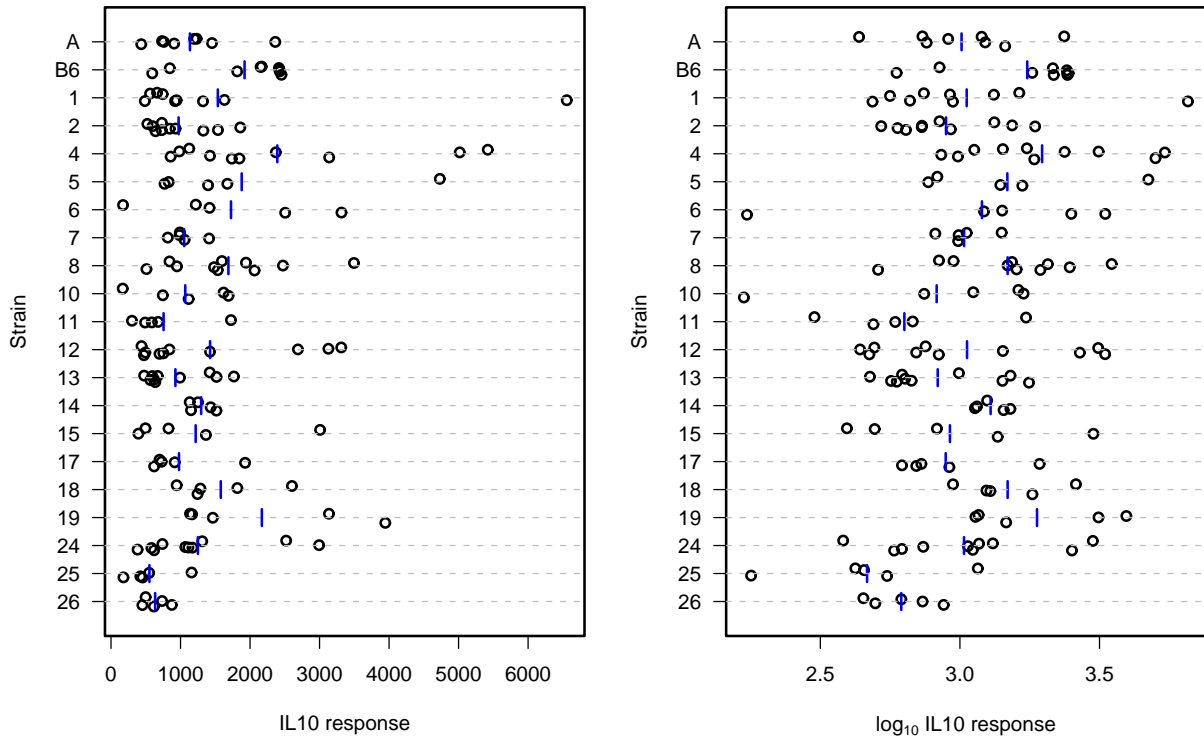
The function `kruskal.test()` in R will calculate the statistic.

Note

In the case of two groups, the Kruskal-Wallis test reduces exactly to the **Wilcoxon rank-sum test**.

This is just like how **ANOVA** is equivalent to the **two-sample t test**.

Example



ANOVA Tables

Original scale / 1000:

source	SS	df	MS	F	P-value
between strains	33	20	1.69	1.70	0.042
within strains	124	125	0.99		
total	157	145			

permutation P-value = 0.043

log₁₀ scale:

source	SS	df	MS	F	P
between strains	3.35	20	0.167	2.25	0.0036
within strains	9.29	125	0.074		
total	12.63	145			

permutation P-value = 0.003

K-W results

The observed Kruskal-Wallis statistic for these data was **41.32**. (Note that it doesn't matter whether you take logs.)

Since there were 21 strains, we can compare this to a χ^2 distribution with **20** degrees of freedom. Thus we obtain the **P-value = 0.003**.

With a permutation test, I got $\hat{P} = 0.0015$ (on the basis of 10,000 simulations).

In the case of ties...

In the case of ties, we assign the **average rank** to each.

Example:

A:	3.5	3.7	4.0	4.2	4.3				
B:			3.9		4.3	4.5			
C:	3.1	3.6		4.0	4.3				
	(1)	(2)	(3)	(4)	(5)	(6/7)	(8)	(9/10/11)	(12)
					↓		↓		
					6.5		10		

Then we apply a correction factor.

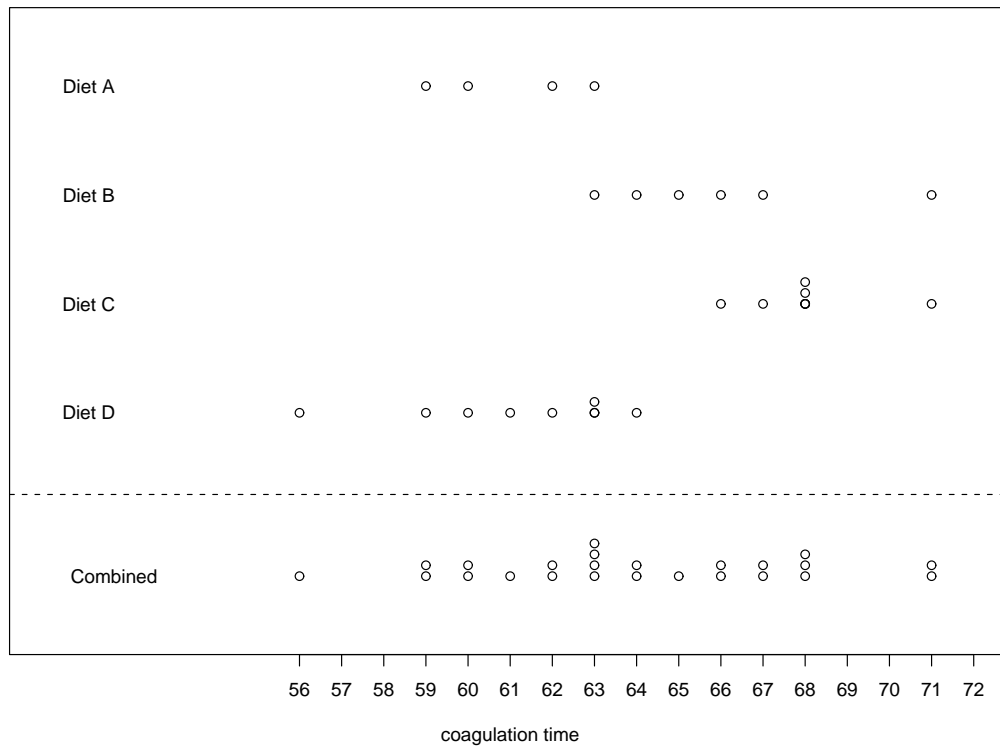
Let $N = \sum_t n_t$ and $T_i =$ no. observations in the i th set of ties (can be 1).

$$\text{Let } D = 1 - \frac{\sum_i (T_i^3 - T_i)}{(N^3 - N)}$$

Use the statistic $H' = H/D$.

Note that $D \leq 1$ and so $H' \geq H$. For the example, $D = 1 - \frac{(2^3-2)+(3^3-3)}{12^3-12} \approx 0.983$.

Blood coagulation time



Example (continued)

A	B	C	D	rank	avg rank
			56	1	1
59				2	2.5
		59		3	2.5
60				4	4.5
		60		5	4.5
		61		6	6
62				7	7.5
		62		8	7.5
63				9	10.5
	63			10	10.5
		63		11	10.5
		63		12	10.5
64				13	13.5
		64		14	13.5
65				15	15
66				16	16.5
	66			17	16.5
67				18	18.5
	67			19	18.5
68				20	21
	68			21	21
	68			22	21
71				23	23.5
	71			24	23.5

Example (continued)

A	62	60	63	59					61	
	7.5	4.5	10.5	2.5					6.25	
B	63	67	71	64	65	66				66
	10.5	18.5	23.5	13.5	15.0	16.5				16.25
C	68	66	71	67	68	68				68
	21.0	16.5	23.5	18.5	21.0	21.0				20.25
D	56	62	60	61	63	64	63	59	61	
	1.0	7.5	4.5	6.0	10.5	13.5	10.5	2.5	7.00	

Calculation of K-W test statistic

	A	B	C	D	
n_t	4	6	6	8	$N = 24$
\bar{R}_t	6.25	16.25	20.25	7.00	$\frac{N+1}{2} = 12.5$

$$\begin{aligned}
 H &= \frac{12}{N(N+1)} \sum_t n_t \left[\bar{R}_t - \left(\frac{N+1}{2} \right) \right]^2 \\
 &= \frac{12}{24 \times 25} \left\{ 4 \times (6.25 - 12.5)^2 + \dots + 8 \times (7.00 - 12.5)^2 \right\} \\
 &= 16.86
 \end{aligned}$$

The ties: $T_i = (1 \ 2 \ 2 \ 1 \ 2 \ 4 \ 2 \ 1 \ 2 \ 2 \ 3 \ 2)$

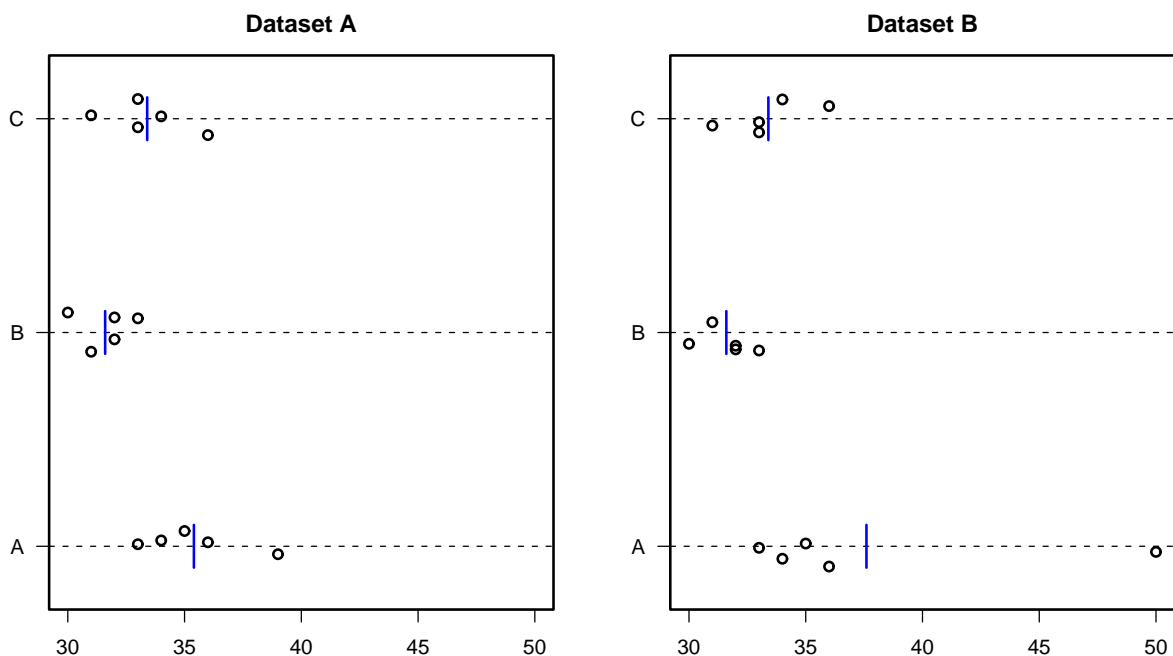
$$D = 1 - \sum_i (T_i^3 - T_i) / (N^3 - N) = \dots = 0.991$$

$$H' = H/D = 16.86 / 0.991 = 17.02 \quad [df = 3] \quad P\text{-value} \approx 0.0007$$

A few points

- **Calculation of P-values:** (avoiding type I errors)
 - F statistic: F distribution (requires normality)
 - K-W statistic: χ^2 distribution (requires large samples)
 - Either statistic: Permutation tests
- **Power:** (avoiding type II errors)
 - K-W statistic more resistant to outliers
 - F statistic more powerful in the case of normality
- K-W statistic: don't need to worry about transformations.

A fake example



Results

Dataset	Method	Statistic	nominal	Permu'n
			P-value	P-value
A	ANOVA	5.48	0.020	0.017
	K-W	7.64	0.022	0.012
B	ANOVA	2.64	0.112	0.023
	K-W	7.64	0.022	0.012

Distributions

