Nonparametric Analysis of Ordinal Data from Designed Experiments

Denis Shah, NYSAES

Larry Madden, OSU

What we will cover

- Ordinal rating scales
- Nonparametric model
- Hypotheses, relative effects, test statistics
- SAS programs and macros

What we will assume you know

- (Some) experimental design
- Some familiarity with SAS (not necessarily with Proc Mixed)

Goals

- Appreciation of what experimental designs can be used if collecting ordinal data
- How to run the analyses
- How to interpret the output
- What to present in your publications

Experimental Design & Data Analysis

Layouts

- 1-way
- 2-way factorial
- Split plot
- Repeated measures

Measurement scales

- Continuous
- Discrete (count)
- Binary (0, 1)
- Ordinal (ordered categories)

How will the data be analyzed?

Common measurement scales

- Continuous (e.g. yield, weight)
- Count (0,1,2,…)
- Proportional/percent (0-1, 0-100%)
- Nominal (numbers serve only to ‘name’ a category)
- Ordinal scale (numerical order has meaning)

Properties of an ordinal scale

- The comparisons between measurements is relevant (>, =, <)
- Numeric values are used only to arrange the measurements from smallest to largest
- Ordering based on relative size

Some nonparametric tests for ordinal (or continuous) data

Type of experimental layout	Test (example)
One random sample	Quantile test
Paired observations	Sign test
Randomized complete block (with single treatment factor)	Freidman ←
Two random samples (groups)	Mann-Whitney ←
Several random samples (but only one factor – not factorial)	Median test Kruskal-Wallis test ←

Rank-based tests

None of these are for factorials, split-plots, etc.

What is a ‘factorial’?

- A class of experiments in which the treatments have a well-defined structure
- Factorial treatments are formed from combinations of two or more different factors
- Each treatment combination must contain one level of every factor

a_1b_1 a_1b_2

a_2b_1 a_2b_2

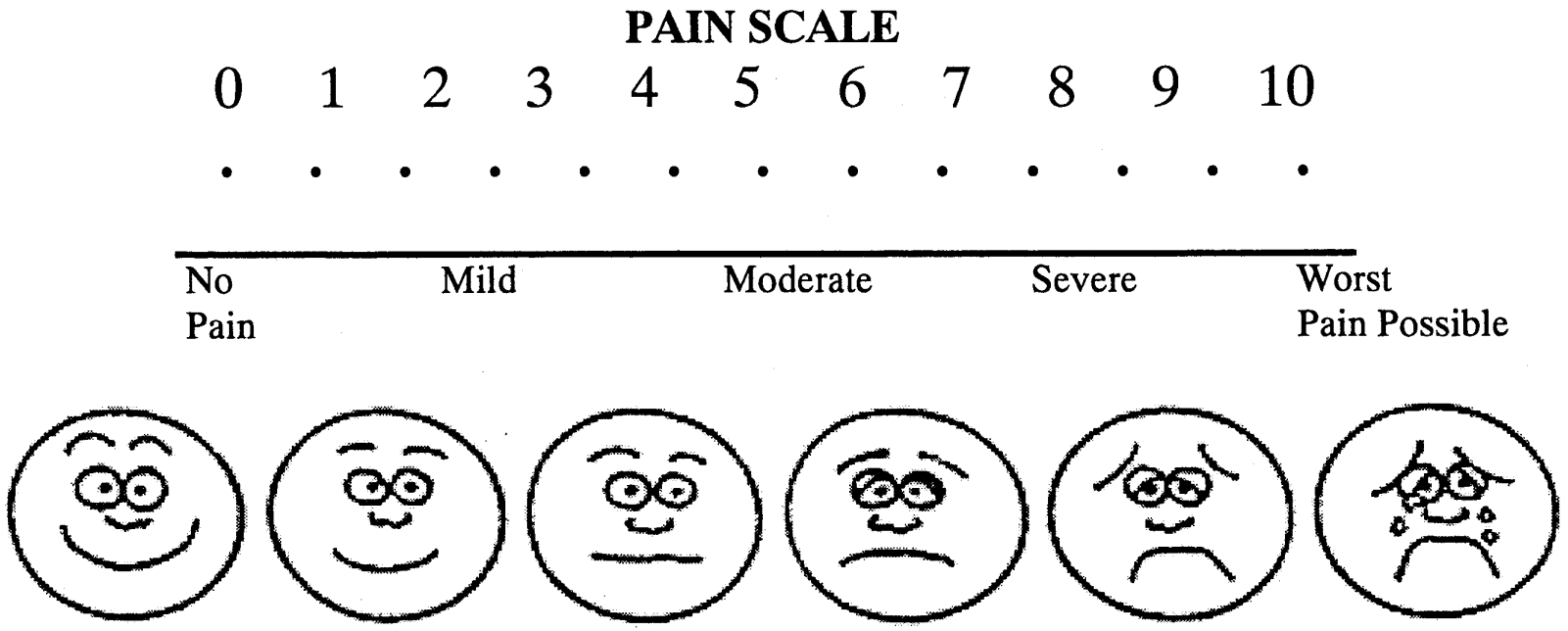
Other nonparametric tests

- Other tests, such as aligned ranks, are available for more complicated designs (multivariate, split plot etc.)
- Generally assume that data are obtained on a continuous scale (i.e. not applicable to ordinal data)

Not covered in this workshop

Examples of ordinal scales

Wong/Baker Faces Pain Scale



Clinical study of multiple sclerosis

Kurtzke Functional Systems Scores (FSS)

Pyramidal Functions

0 - Normal
1 - Abnormal signs without disability
2 - Minimal disability
3 - Mild to moderate paraparesis or hemiparesis (detectable weakness but most function sustained for short periods, fatigue a problem); severe monoparesis (almost no function)
4 - Marked paraparesis or hemiparesis (function is difficult), moderate quadriparesis (function is decreased but can be sustained for short periods); or monoplegia
5 - Paraplegia, hemiplegia, or marked quadriparesis
6 - Quadriplegia
9 - (Unknown)

Hauser Ambulation Index

- 0 = Asymptomatic; fully active.
- 1 = Walks normally, but reports fatigue that interferes with athletic or other demanding activities.
- 2 = Abnormal gait or episodic imbalance; gait disorder is noticed by family and friends; able to walk 25 feet (8 meters) in 10 seconds or less.
- 3 = Walks independently; able to walk 25 feet in 20 seconds or less.
- 4 = Requires unilateral support (cane or single crutch) to walk; walks 25 feet in 20 seconds or less.
- 5 = Requires bilateral support (canes, crutches, or walker) and walks 25 feet in 25 seconds or less; *or* requires unilateral support but needs more than 20 seconds to walk 25 feet.
- 6 = Requires bilateral support and more than 20 seconds to walk 25 feet; may use wheelchair* on occasion.
- 7 = Walking limited to several steps with bilateral support; unable to walk 25 feet; may use wheelchair* for most activities.
- 8 = Restricted to wheelchair; able to transfer self independently.
- 9 = Restricted to wheelchair; unable to transfer self independently.

Ordinal rating scales are common in plant pathology

- Root diseases
- Foliar diseases
- Diseases of fruit, berries etc…..

Fusarium root-rot severity

1 No visible symptoms

3 One to 3 leaves, representing no more than 10% of the total foliage, are wilted and chlorotic

5 Approximately 25% of leaves and branches exhibit wilting and chlorosis

7 Approximately 50% of leaves and branches exhibit wilting and chlorosis

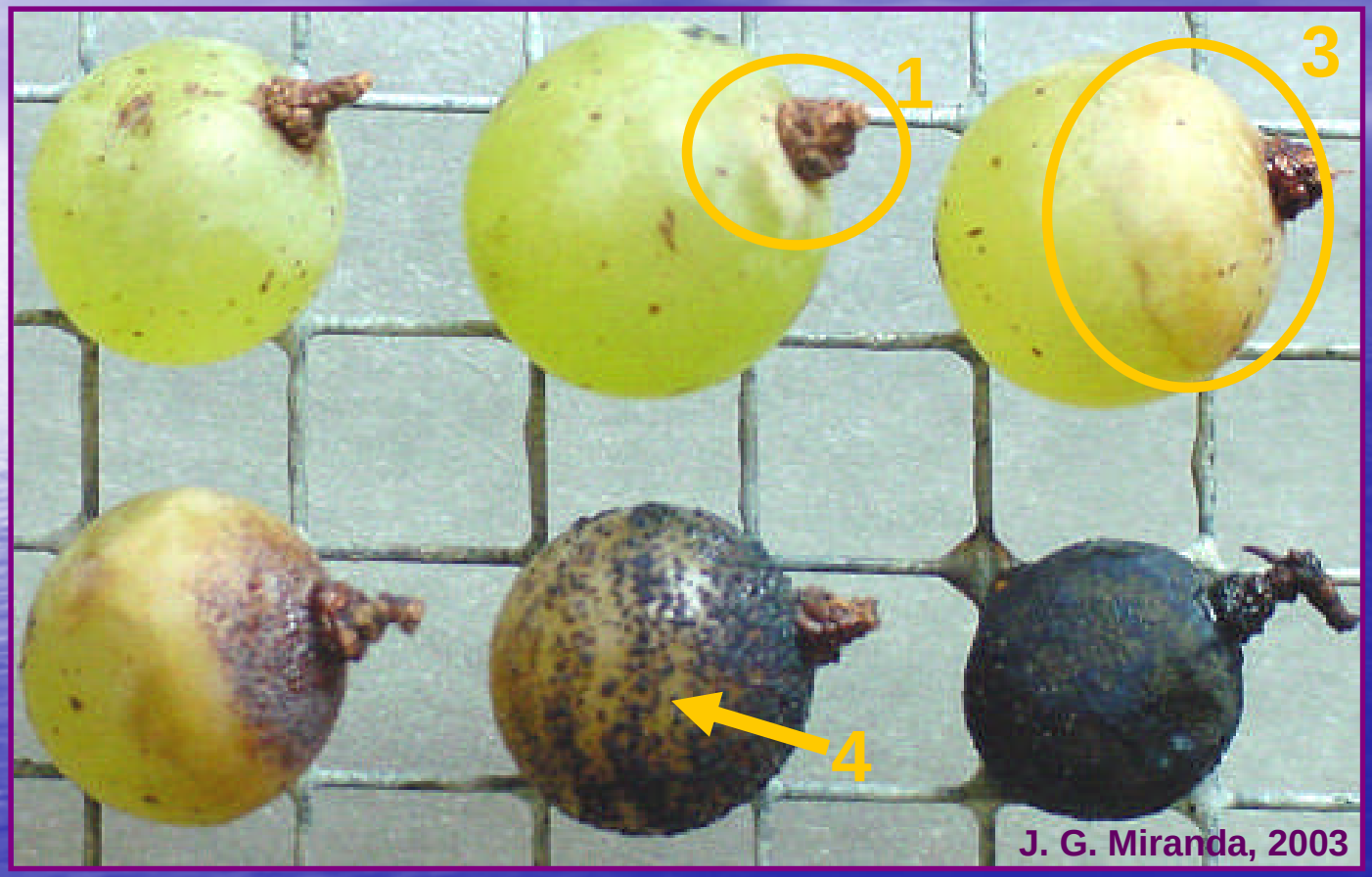
9 Approximately 75% or more of the leaves and branches exhibit wilting, chlorosis, and defoliation, with eventually plant death

Ceballos et al. 2004. Effect of five postemergence herbicides on red clover shoot and root growth in greenhouse studies. Phytoprotection 85:153-160.

- Root injury
  - 1 = no symptoms
  - 2 = lesions present
  - 3 = necrosis
- Shoot phytotoxicity
  - 1 = no visible damage
  - …
  - 5 = plant is dead

Carrot Incidence and Severity of RKN Infection

Symptom development of bitter rot



Disease Rating Scale	
0	No infection
1	1 – 5 % infection
2	6 – 15 % infection
3	16 – 50 % infection
4	≥ 51% infection

Russet on snap bean pods

- 0 = no symptoms
- 1 = a few flecks
- 2 = 2-5% of pod covered
- 3 = 5-10%
- 4 = 10-25%
- 5 = 25-50%
- 6 =50-70%
- 7 = 70-90%
- 8 = 90-<100%
- 9 = 100%

Stagonospora nodorum leaf blotch of wheat.

Liu et al. 2004. Phytopathology 94: 1061-1067.

Horsfall-Barratt Scale

0 to 11 rating scale representing percent disease severity

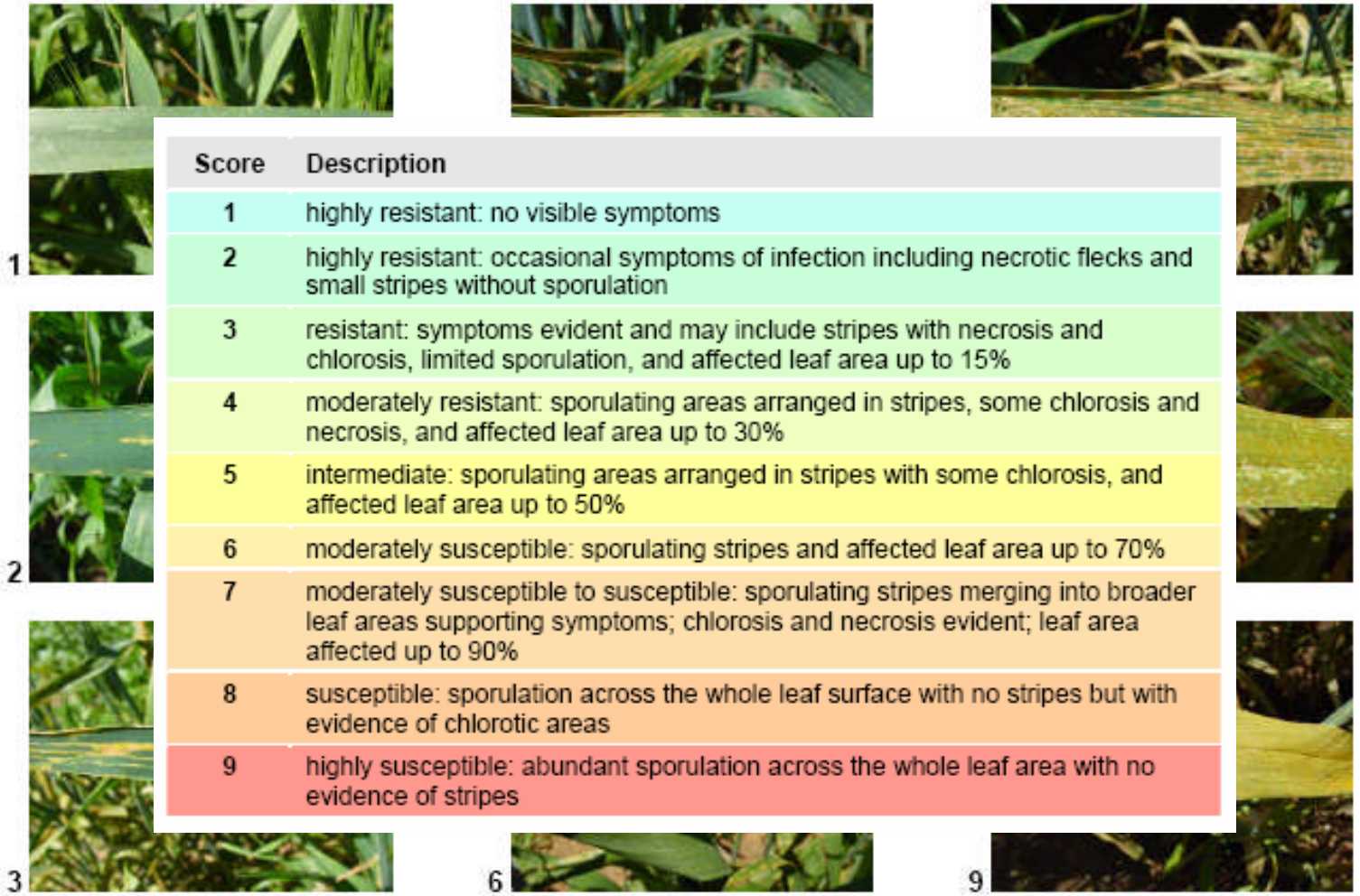
0 = no disease	1 = 1-3%	2 = 4-6%
3 = 7-12%	4 = 13-24%	5 = 25-50%
6 = 51-75%	7 = 76-88%	8 = 89-94%
9 = 95-97%	10 = 98-99%	11 = 100%

Barratt, R.W. and J. G. Horsfall. 1945. An Improved Grading System for Measuring Plant Disease. Connecticut Agricultural Experiment Station.

Carrot Forecasting Trial - Hancock, 2002-03

- 2 cultivars: Bolero & Fontana
- 4 Treatments: chlorothalonil (1.2 lb ai/A) alt. azoxystrobin (0.15 lb ai/A)
- Treatment initiation at 1 % severity threshold
- Foliar disease severity (%) rated every 7 days on H-B scale (0-11)
- Treatments evaluation:
  - Disease severity (weekly)
  - AUDPC (season)
  - Yield, quality and value

Stripe rust on wheat



Score	Description
1	highly resistant: no visible symptoms
2	highly resistant: occasional symptoms of infection including necrotic flecks and small stripes without sporulation
3	resistant: symptoms evident and may include stripes with necrosis and chlorosis, limited sporulation, and affected leaf area up to 15%
4	moderately resistant: sporulating areas arranged in stripes, some chlorosis and necrosis, and affected leaf area up to 30%
5	intermediate: sporulating areas arranged in stripes with some chlorosis, and affected leaf area up to 50%
6	moderately susceptible: sporulating stripes and affected leaf area up to 70%
7	moderately susceptible to susceptible: sporulating stripes merging into broader leaf areas supporting symptoms; chlorosis and necrosis evident; leaf area affected up to 90%
8	susceptible: sporulation across the whole leaf surface with no stripes but with evidence of chlorotic areas
9	highly susceptible: abundant sporulation across the whole leaf area with no evidence of stripes

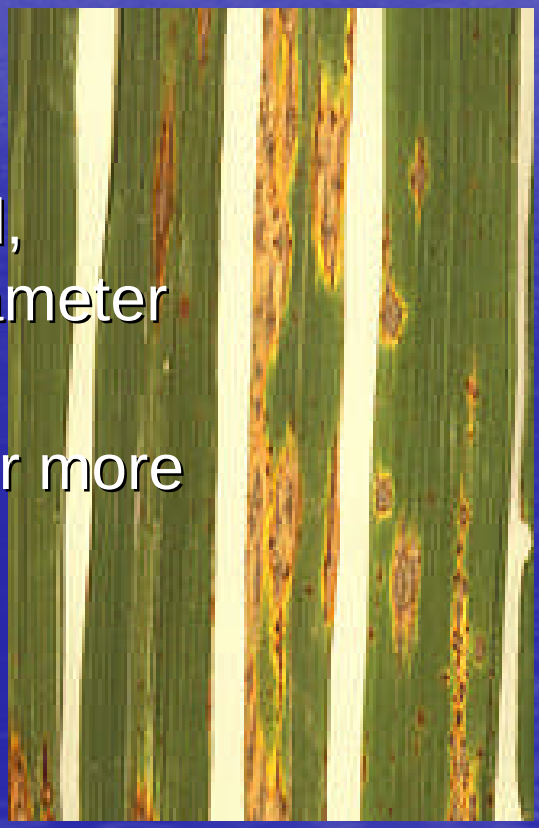
Table 2. Results of new product testing for control of *Cercospora* and *Alternaria* blights of carrot.

Treatment and rate/A (application sequence[z])	Petiole blight				Petiole health[w]		Leaf blight (%)[v]		Yield per 10-ft row (lb)	
	Incidence (%)[y]		Severity[x]							
Untreated	97.9	e[u]	5.0	c	7.5	d	68.2	d	14.7	c
Bravo Weather Stik 6SC 1.5 pt (1-10)	22.0	abcd	1.5	ab	1.8	a b	5.5	a b c	24.6	a b
Bravo Weather Stik 6SC 1.5 pt (1,3,5,7,9)	74.6	de	2.5	b	3.3	b c	13.8	c	25.1	a b
Bravo Weather Stik 6SC 1 pt (1-10)	18.4	abc	1.8	ab	2.0	a b c	4.2	a b	25.0	a b

x Petiole blight severity rated on a 1 to 5 scale; where 1 = 0 petiole lesions per plant, 2 = 1-10, 3 = 11-21, 4 = 21-50, and 5 = > 50.
w Petiole health rated on a 1 to 10 scale; where 1 = healthy and vigorous to 10 = necrotic or dead.

Mackill & Bonman. 1986. New hosts of *Pyricularia oryzae*. Plant Dis. 70: 125-127.

- 0 = no infection
- 1 = small brown specks of pinhead size
- 2 = 1.5 mm brown specks
- 3 = small, roundish to slightly elongated, necrotic gray spots about 2-3 mm in diameter with brown margins
- 4 = typical blast lesions infecting 50% or more of the leaf area



Bosland & Lindsey. 1991. A seedling screen for Phytophthora root rot of pepper, *Capsicum annuum*. Plant Dis. 75: 1048-1050.
(0-10 scale)

- 0 = no response
- 3 = brown roots, slight stunting, very small lesions on stems
- 7 = brown roots, large lesions on stems, girdling, whole plant wilted, and stunted
- 10 = death

Van Toai et al. 1994. Genetic variability for flooding tolerance in soybeans. Crop Sci. 34:1112-1115.

- 1 = healthy plants with no root root
- …
- 10 = all seedlings killed

Pratt et al. 1994. Maize responses to a severe isolate of maize chlorotic dwarf virus. Crop Sci. 34:635-641.

Chlorosis

- 1 = no symptoms
- 2 = Chlorosis just beginning
- 3 = Chlorosis is clearly visible in base of two youngest leaves
- 4 = In addition to 3, chlorosis on at least one-half the length of three to four youngest leaves
- 5 = Chlorosis more severe than in 4, leaves are yellow and are beginning to turn white

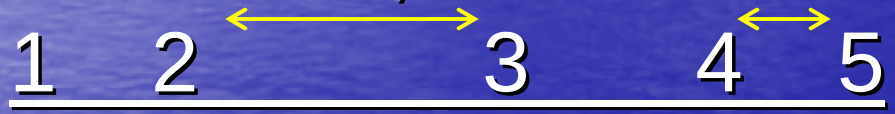
U. Of Georgia Cooperative Extension Guide

Dealing with ordinal data

- Differences between scores (or mean scores) do not make sense
- Therefore, methods based on the analysis of means (ANOVA) are not appropriate
- The results should not depend on the values assigned to the categories (the 'labels'). i.e. the results should be invariant (same) under monotonic transformations of the rating scale. Analysis based on rank transformations can meet these criteria.

Difference between scores do not make sense (in any quantitative or physical way)

For all we know, the scale could look like this:



Or this: 

Or even this: 

Defining ranks

E.g., 2 treatments, effect measured on a 0-4 ordinal scale

Trt 1	Trt 2
1	3
0	2
1	4

Trt1	Trt2
2.5	5.0
1.0	4.0
2.5	6.0

Go to SAS…

Rank-based tests

- Have been around for a long time (Kruskal-Wallis, Friedman)
- But generally limited to the one-way layout (i.e., there *had* been sound statistical theory for ordinal data only for the one-way layout)
- Given the desirable properties of rank transformations, why not use ANOVA on the ranks (i.e. Rank Transform Method)?

Don’t just use ANOVA on ranks!

- Hypotheses in ANOVA are based on differences between means, or shifts in means (“expected values”). These are affected by monotonic data transformations. Rank statistics are invariant, so inappropriate to use them to test hypotheses that are transformation-dependent.
  - Looked at another way, if one uses ranks of data, one is not testing the equality of means (expected values) for different treatments

Don’t just use ANOVA on ranks!

- Assumption of normality in classical ANOVA: ranks are not normally distributed
- Ranked data have unequal variances, even if the variances were constant in the original data

Getting around ordinal data: the disease index

- A common approach in plant pathology

Example: Kora et al. CJPP 2005

0 = 0%

1 = 1-25%

2 = 26-50%

3 = 51-75%

4 = 76-100%

$$D_{index} = \frac{\sum(\text{severity class} \times \text{no. roots in class})}{(\text{total roots} \times \text{highest class No.})} \times 100$$

It is debatable if such an approach is justified.

Another example

"Roots were washed and evaluated for disease using a 0 to 4 rating scale. A disease severity index (DSI) was calculated for each plot by: (mean severity X incidence %) / 4. "

Bradley et al. (web document)

Statistical issues

A new approach (subject of this workshop)

- Applicable to continuous, discrete, dichotomous or ordinal data
- Robust with respect to outliers
- Results are invariant under strictly monotone transformations of the data
- Missing values are allowable
- Very good approximate test statistics are available for small sample sizes

A new approach (M. Akritas, Edgar Brunner & several colleagues)

- Most (routine) experimental designs (layouts) can be handled with specialized, free macros (SAS or R)
- Designs (plus contrasts) can be generally handled with SAS Proc Mixed (with appropriate options)

Assumptions

- Nonparametric does not mean there are no assumptions
  - All statistical methods are based on assumptions
- The Brunner approach has the least restrictive assumptions of all possible statistical methods for testing hypotheses about random variables

Assumptions

- Other nonparametric tests have more restrictive assumptions:

  E.g., K-W (which is strictly for a one-way layout) assumes:

  – constant variance across groups: $S^2 = N(N+1)/12$ when there are no ties (i.e., for continuous data).

  – Distributions of observations have the same shape for all groups (treatments, etc.), when one is testing for equality of medians

- K-W can be regarded as a special case of the Brunner one-way layout.

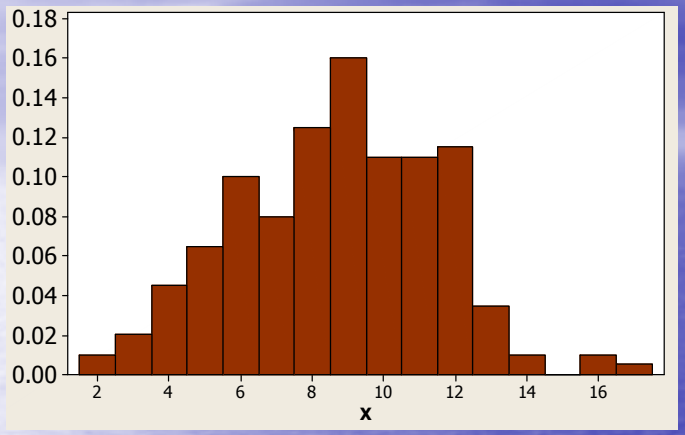
Assumptions in the Brunner approach

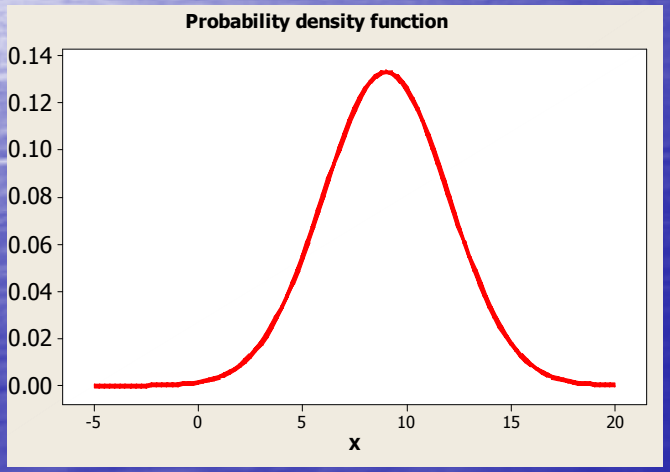
- Observations have a distribution!
  - (no restrictions on shape of distributions, nor on similarity of distributions among groups)
- There are sufficient number of observations (replications) to apply certain test statistics.
  - In fact, simulations show that the approach works for small sample sizes
- Essentially, no other assumptions.

Nonparametric statistical analysis

- Approach depends on *normalized distributions*, and so-called *relative treatment effects*
- Thus, a little review is provided….

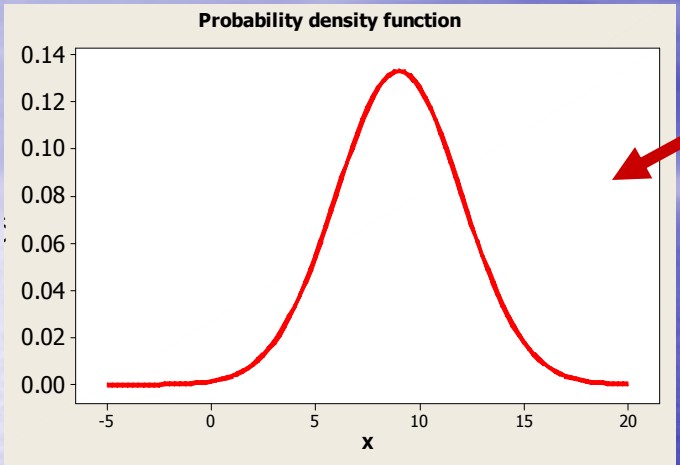
Distributions

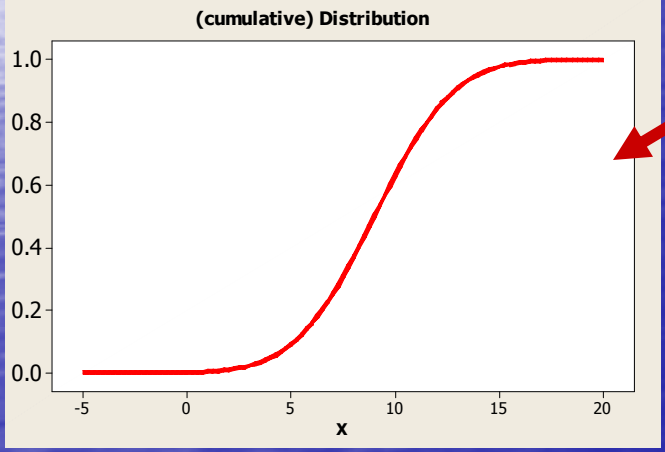




- **Histogram**
  - Division of a sample of observations of a random variable into a number of classes, together with the number (or proportion) of observations in each class
- **Probability density function (*pdf*) or probability mass function (*pmf*)**
  - The probability of each value of a variable in a population (discrete)
  - Probability that a variable falls within a particular interval in a population when integrated over interval (continuous)
  - Sometimes just called the 'distribution' (but not here)
- **Estimated probability density function**
  - Estimated *pdf* from a sample
  - Often called *empirical* probability density
  - Equivalent (graphically) to scaled histogram

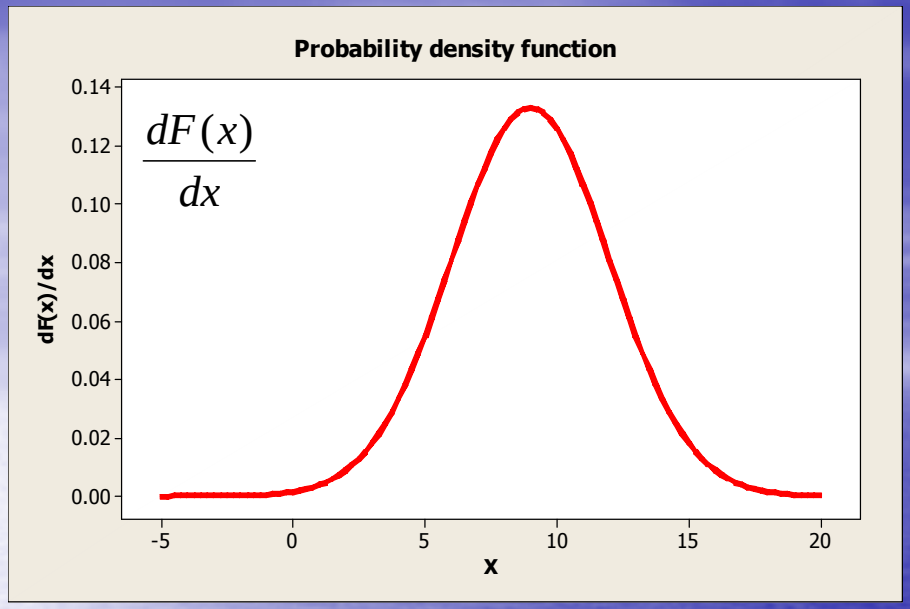
Distributions

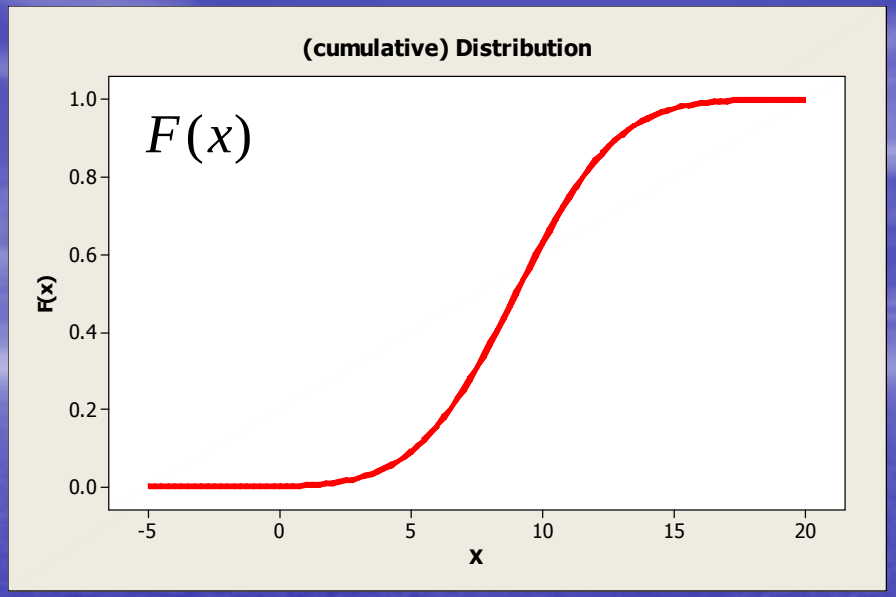


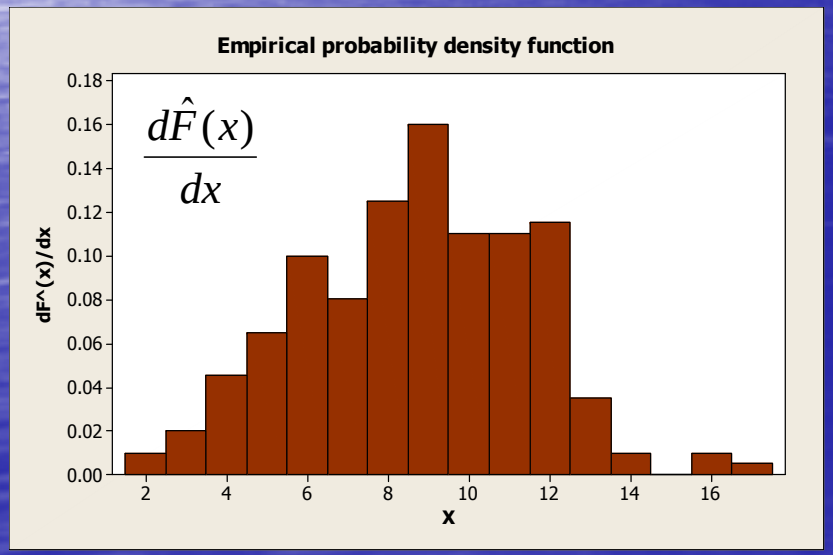


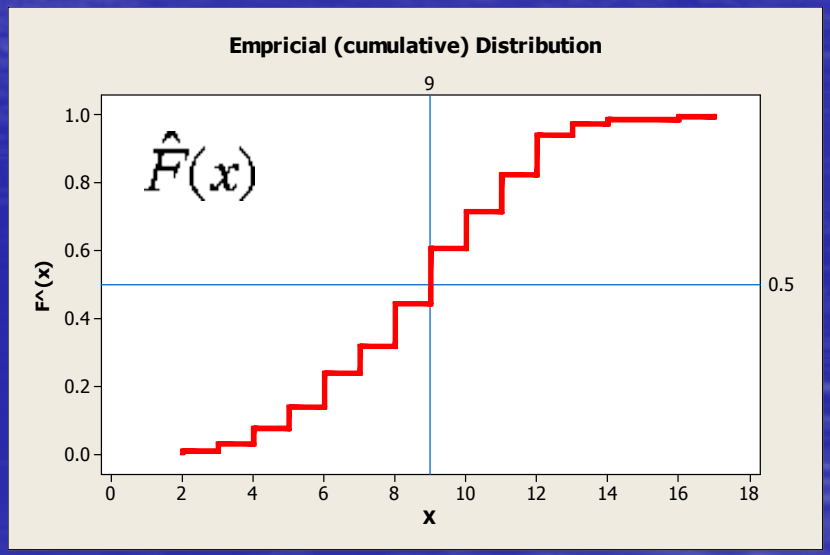
With *F(x)* for distribution, *pdf* is thus written as: *dF(x)/dx*

- **Probability density function (*pdf*) or probability mass function (*pmf*)**
  - The probability of each value of a variable in a population (discrete)
  - Probability that a variable falls within a particular interval in a population (continuous), when integrated over interval
- **Distribution**
  - *Cumulative* probability of values of a variable in a population
    - Labeled as *F*(*x*) or simply *F*
  - Sometimes called cumulative distribution
- **Estimated distribution**
  - Sometimes called empirical distribution
    - Labeled as $\hat{F}(x) = \hat{F}$

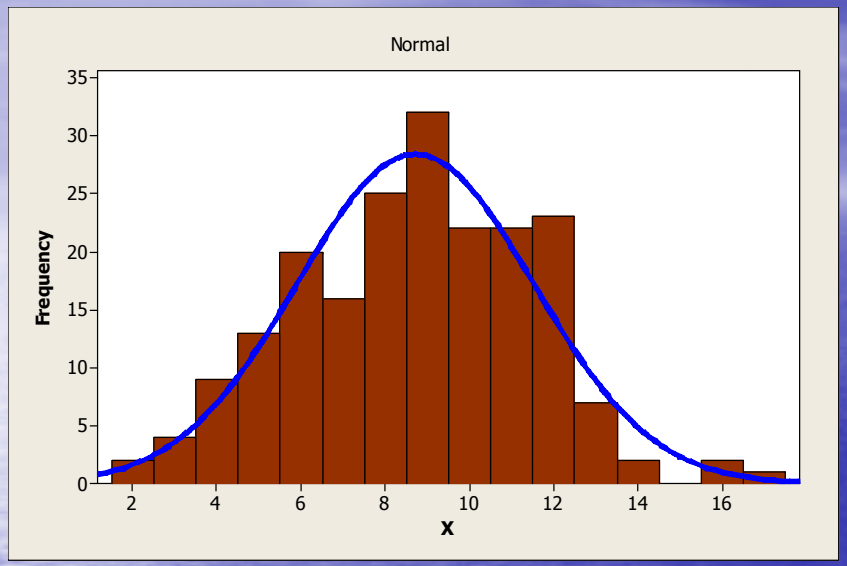

Probability density function
$\frac{dF(x)}{dx}$
dF(x)/dx
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
-5
0
5
10
15
20
X

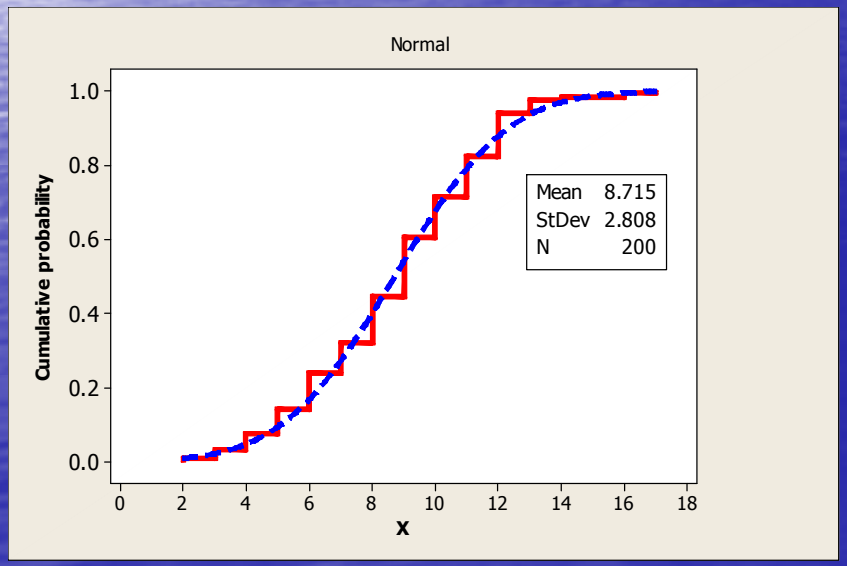

(cumulative) Distribution
$F(x)$
F(x)
1.0
0.8
0.6
0.4
0.2
0.0
-5
0
5
10
15
20
X


Empirical probability density function
$\frac{d\hat{F}(x)}{dx}$
dF^(x)/dx
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
2
4
6
8
10
12
14
16
X


Empricial (cumulative) Distribution
9
$\hat{F}(x)$
F^(x)
1.0
0.8
0.6
0.4
0.2
0.0
0.5
0
2
4
6
8
10
12
14
16
18
X

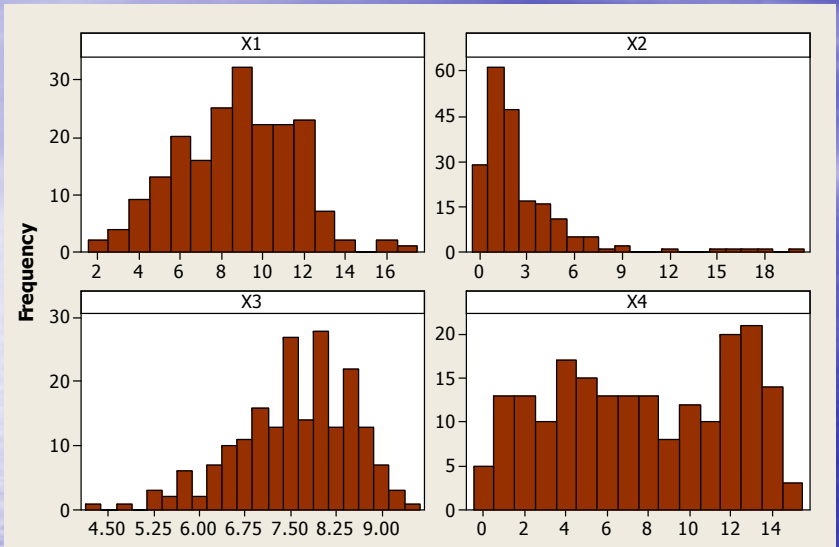
Distributions

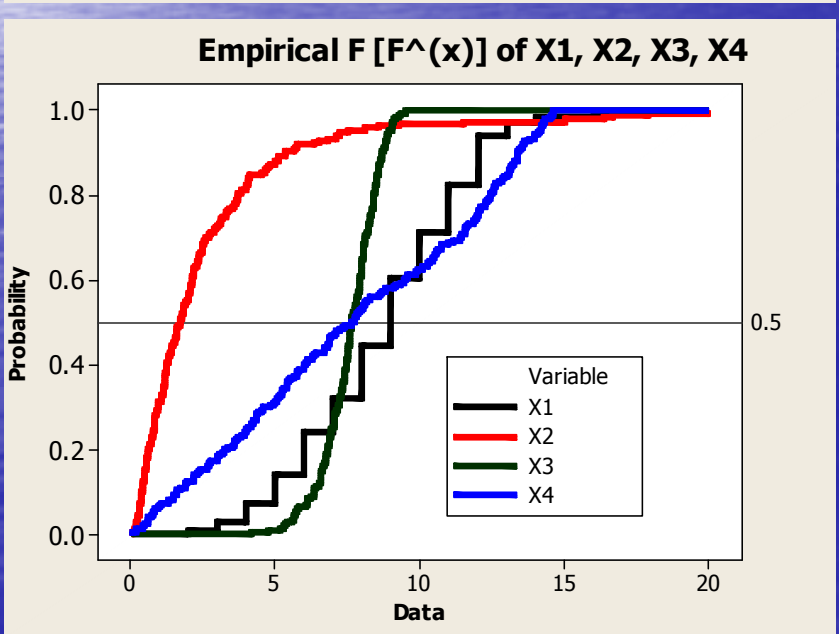




- The foundation of **parametric statistical analysis** is that the distribution (*F*) of a variable can be represented by a function (i.e., model) with one or more parameters
  - Normal distribution
    - Mean (μ)
    - Variance (σ^2)
  - Exponential, gamma, log-normal, Poisson, negative binomial, etc. …
- Descriptions, comparisons, predictions, and in general, inference, are performed in terms of estimated parameters
- With ordinal data, however, this is *not* possible.

Distributions





- In fully **nonparametric statistical analysis,** one does not (generally) assume any function (model) for *F* or *dF/dx*
  - The measurement scale (i.e., type of random variable) precludes use of functions such as the normal, Poisson, and other models for *F*.
    - **Ordinal data**
  - Conditions or assumptions needed (desired) to use certain functions for *F* are violated
- However, with nonparametric statistics, one can base analyses *directly* on distributions and their estimates
  - Basis for this workshop…..

Distributions

- **It turns out that investigators do not actually have to estimate *F*s explicitly**
- However, since the principles and concepts are based on *F*s, it is worth spending a little time working through some calculations for a small data set
  - The calculations lead to a useful summary statistic that is used in the nonparametric analyses of this workshop.
- Consider the following 10 points, for a single group (e.g., treatment)
    - **X_k = 1, 2, 2, 4 , 5, 6, 7, 7, 9, 10 (n = 10; k = 1, 2, …, n)**
  - **What is the empirical (estimated) *F(x)*?**
- Note: Upper case *X* for the random variable, and lower case *x* for a specific (fixed) value
- So far, we have deliberately been a little vague about the cumulative aspect of the probability.
  - The "usual" or "classical" definition is: **Prob[$X \leq x$]**
    - Example: Probability that an observation is less than or equal to *x*=1, 2, …
  - However, there are actually three versions of the distribution.

Distributions

- Before calculating the distribution, first consider the ranks of the observations
- **Ranks**:
  - The relative positions of observations in a sample with respect to some characteristic (e.g., some measurement)
  - Representation of the underlying order of the values of a sample

X	Mid-rank, *R*
1	1
2	2.5
2	2.5
4	4
5	5
6	6
7	7.5
7	7.5
9	9
10	10

There are different types of ranks, but the methods that follow are based completely on **mid-ranks** (***R***)

With mid-ranks, ties have the same value

For simplicity, we refer to mid-ranks as *ranks*

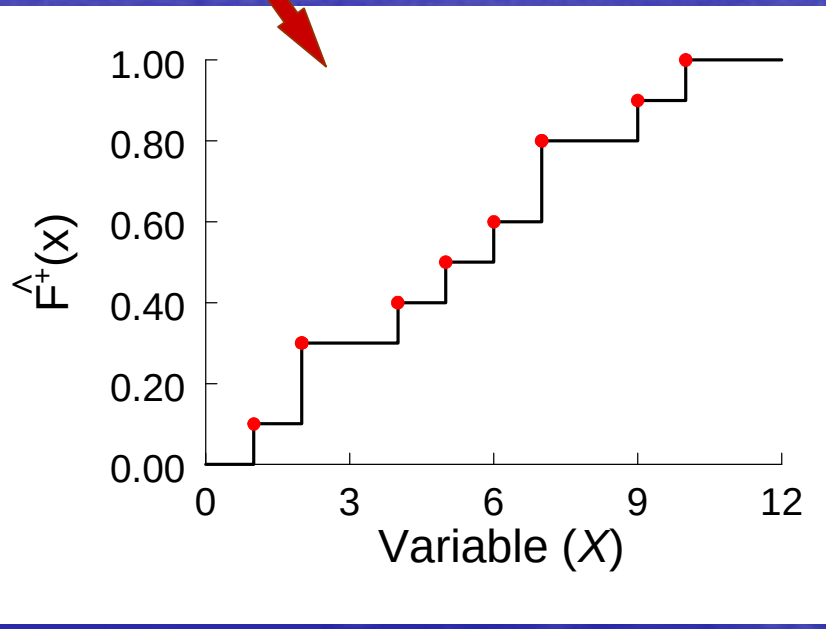
When needed for clarity, use *k* subscript to indicate the specific observation (*k* = 1, …*n*): X_k, R_k

Distributions: Three versions

Right Continuous, $F^{+}(x)$ = Prob[$X \leq x$]

X	R			Prob[$X \leq x$]
1	1			1/10 = 0.1
2	2.5			3/10 = 0.3
2	2.5			3/10 = 0.3
4	4			4/10 = 0.4
5	5			5/10 = 0.5
6	6			6/10 = 0.6
7	7.5			8/10 = 0.8
7	7.5			8/10 = 0.8
9	9			9/10 = 0.9
10	10			10/10=1.0

$\hat{F}^{+}(x)$ Empirical distribution



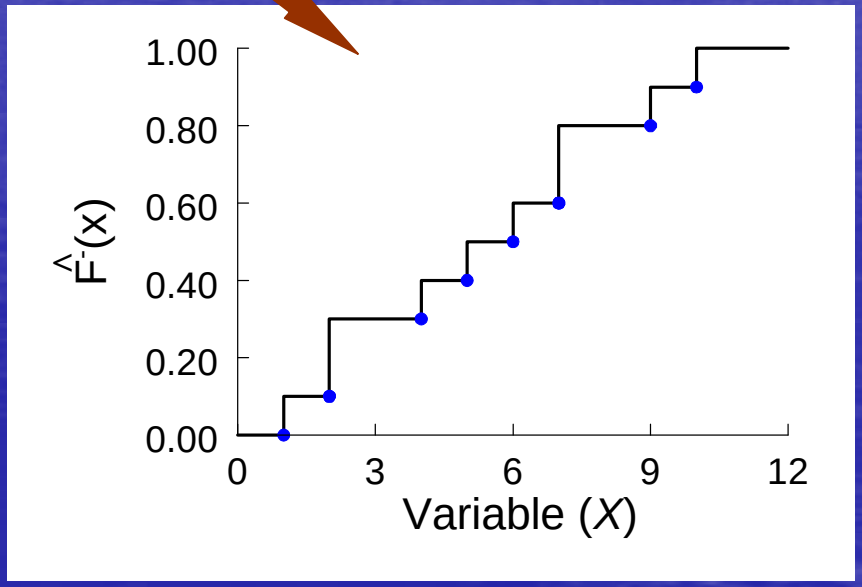
Distributions: Three versions

Left Continuous, $F^{-}(x)$ = Prob[$X<x$]

$\hat{F}^{-}(x)$

Empirical distribution

X	R	Prob[$X < x$]		Prob[$X \leq x$]
1	1	0/10 = 0.0		1/10 = 0.1
2	2.5	1/10 = 0.1		3/10 = 0.3
2	2.5	1/10 = 0.1		3/10 = 0.3
4	4	3/10 = 0.3		4/10 = 0.4
5	5	4/10 = 0.4		5/10 = 0.5
6	6	5/10 = 0.5		6/10 = 0.6
7	7.5	6/10 = 0.6		8/10 = 0.8
7	7.5	6/10 = 0.6		8/10 = 0.8
9	9	8/10 = 0.8		9/10 = 0.9
10	10	9/10 = 0.9		10/10=1.0



Distributions: Three versions

Normalized,

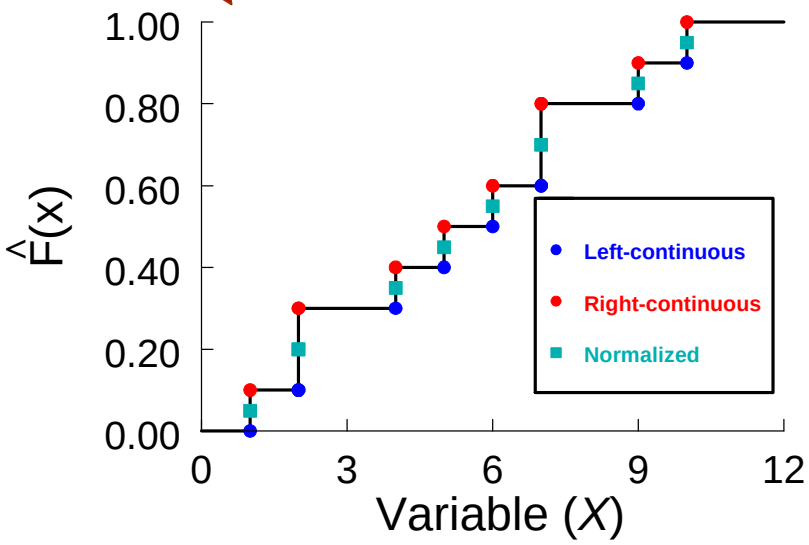
$$F(x) = 0.5 \cdot \{\text{Prob}[X<x] + \text{Prob}[X \leq x]\} = 0.5 \cdot \{F^{-}(x) + F^{+}(x)\}$$

Prob[$X<x$] + 0.5·Prob[$X=x$]

$\hat{F}(x)$

Empirical distribution

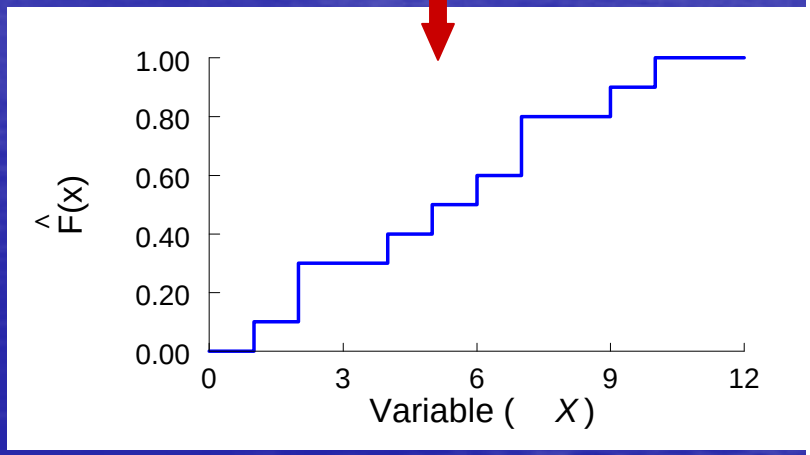
X	R	Prob[$X < x$]	Prob[$X < x$] + 0.5Prob[$X=x$]	Prob[$X \leq x$]
1	1	0/10 = 0.0	0.05	1/10 = 0.1
2	2.5	1/10 = 0.1	0.2	3/10 = 0.3
2	2.5	1/10 = 0.1	0.2	3/10 = 0.3
4	4	3/10 = 0.3	0.35	4/10 = 0.4
5	5	4/10 = 0.4	0.45	5/10 = 0.5
6	6	5/10 = 0.5	0.55	6/10 = 0.6
7	7.5	6/10 = 0.6	0.7	8/10 = 0.8
7	7.5	6/10 = 0.6	0.7	8/10 = 0.8
9	9	8/10 = 0.8	0.85	9/10 = 0.9
10	10	9/10 = 0.9	0.95	10/10=1.0



Distributions: summary (so far)

- *F(x)* or *F* represents the normalized distribution
- Estimated (empirical) normalized distribution indicated with a "hat"
- Density (*pdf*), and hence histogram, is given by *dF/dx*
- *F* gives a full description of the observations
- In nonparametric analysis, no assumptions are needed about the nature of *F*
  - Variable can be continuous or discrete, including ordinal and categorical
  - Ties are permitted

$\hat{F}(x) \leftrightarrow \hat{F}$



Distributions: several groups

- What if there are several groups (treatments)?
- Place a subscript on *F* to indicate the group
  - **F_1, F_2, …., F_A for *a* different groups**
  - Use *i* as a label for a specific group
    - **F_i, $i = 1,\ldots a$**
  - The random variable and rank now have two subscripts, **X_{ik}** and **R_{ik}** (for group and observation)
- One can, if desired, estimate *F* for each group (i.e., determine the empirical distribution for each) $\Leftrightarrow \hat{F}_i$
  - Analysis does not require explicit estimation of F_i.
- A weighted mean *F* (= *H*) can be determined

$$H(x) = H = \frac{1}{N}\sum_1^a n_i F_i$$

Total observations

Observations in group *i*

One can determine empirical *H* ($\hat{H}$) based on empirical *F*

Example empirical normalized distributions

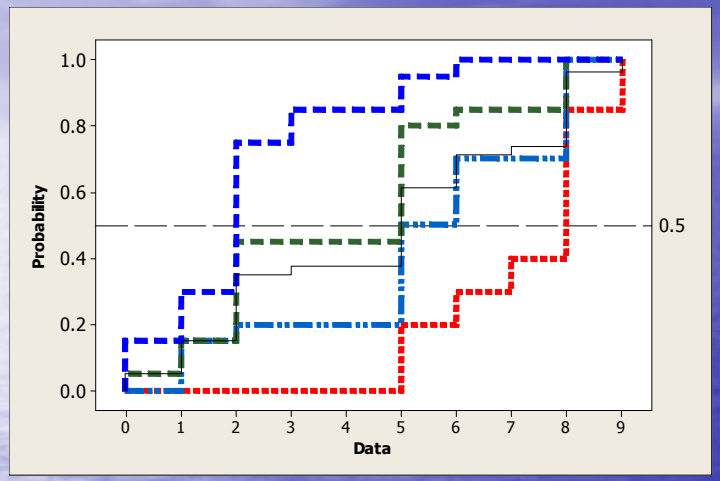
Go to SAS file for example of wheat powdery mildew

(4 cultivars, 20 plants each)

Rating scale

0	<1% leaf area affected on 4th leaf^{a}
1	1-50% leaf area affected on 4th leaf
2	1-5% leaf area affected on 3rd leaf
3	5-15% leaf area affected on 3rd leaf
4	>15% leaf area affected on 3rd leaf
5	1-5% leaf area affected on 2nd leaf
6	5-15% leaf area affected on 2nd leaf
7	>15% leaf area affected on 2nd leaf
8	1-5% leaf area affected on flag leaf
9	5-15% leaf area affected on flag leaf
10	>15% leaf area affected on flag leaf

Normalized distributions: comparisons



- Need a summary value for each distribution to facilitate comparisons of distributions
  - **Are the values of *X* for one group larger (smaller) than for another group?**
- As indicated before, there is no parameter to compare for nonparametric analyses
- The **median** is a useful summary statistic, corresponding to the value of *X* giving ***F*(x) = 0.5**.
  - Some nonparametric approaches are based on medians
  - However, these approaches are not applicable for factorials (repeated measures, etc.), but medians are still useful summaries

Effects of treatments (cultivars, controls, pathogen races, etc.) are defined and determined based on distributions

Relative treatment effects

- A more informative and useful metric than the median is the **relative treatment effect** (also known as the **relative marginal effect** for factorials)
  - $p_i = \int H dF_i$
  - A quantity to represent **the probability that one random variable is larger than the other**
  - Range: $\mathbf{0 < p_i < 1}$ (not quite 0 or 1 for the limits)
- Formally, **p_i quantifies the (stochastic) *tendency* of the distribution F_i with respect to the mean distribution H**
  - If F_i tends to lie to the right of H, then $p_i > 0.5$
  - If F_i tends to lie to the left of H, then $p_i < 0.5$
  - Describes how the observations of one group (with distribution F_i) are related to observations from a group with distribution H
    - If $p_i < 0.5$, there is a tendency of randomly selected observations from group i to be smaller than randomly selected observations from a hypothetical group with H as its distribution

Relative treatment effects

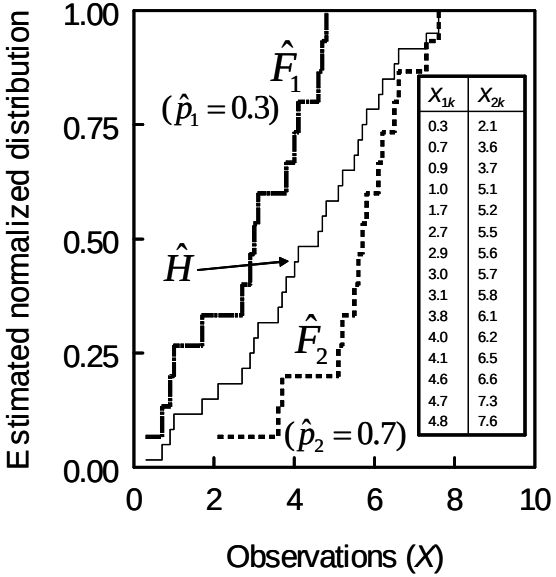
- **Relative treatment effect**: $p_i = \int H dF_i$
- **Estimate**:

Definition of estimate, but not the practical approach for estimation

$$\hat{p}_i = \int \hat{H} d\hat{F}_i$$

$$= \frac{\left(\overline{R}_{i\bullet} - 0.5\right)}{N}$$

It turns out that the estimate is a simple function of the mean rank for the *i*-th group



$Med_1 = 3.0$ $\quad$ $Med_2 = 5.7$

$\overline{R}_{1\bullet} = 9.5$ $\quad$ $\overline{R}_{2\bullet} = 21.5$

$\hat{p}_1 = \frac{9.5 - 0.5}{30} = 0.3$ $\quad$ $\hat{p}_2 = \frac{21.5 - 0.5}{30} = 0.7$

Reminder:

R_{ik}:	Rank of *k*-th observation in group *i*
N:	Total number of observations
H:	Weighted mean normalized distribution
dF/dx:	Probability density function

Relative treatment effects

- When there are only two groups, one can define:

  $$p = p_2 - p_1 + 0.5 = \text{Prob}(X_1 < X_2) + 0.5 \cdot \text{Prob}(X_1 = X_2) = \int F_1 dF_2$$

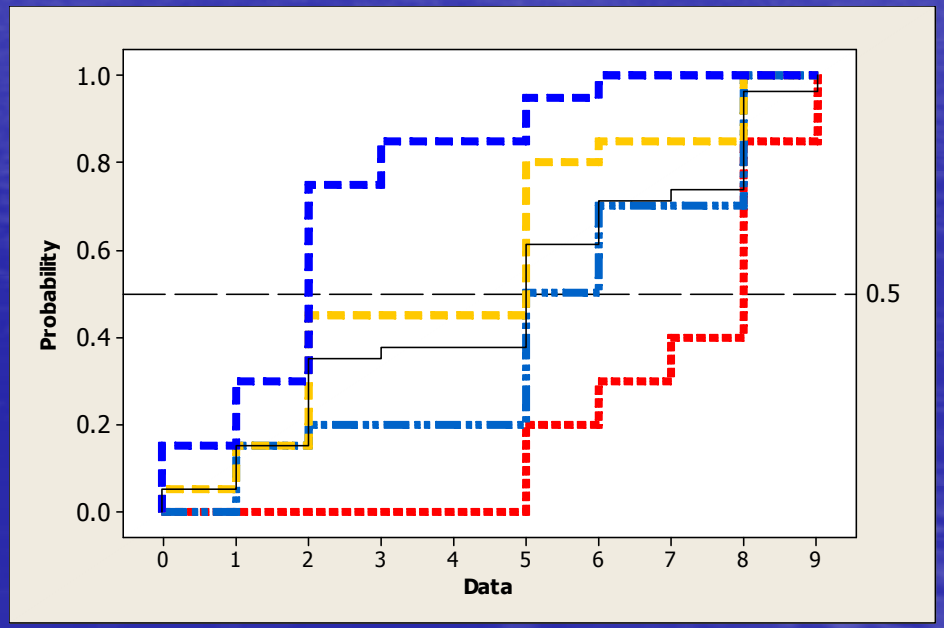
  - ***The relative effect of F_2 with respect to F_1***
  - **"The probability that the random variable from group 2 is greater than from group 1"**
    - **$p > 0.5$ ($p_2 - p_1 > 0$):** Values of X_2 tend to be larger than values of X_1
    - **$p < 0.5$ ($p_2 - p_1 < 0$):** Values of X_2 tend to be smaller than values of X_1
    - **$p = 0.5$ ($p_2 - p_1 = 0$):** No tendency exists for the values of X_1 to be either larger or smaller than those of X_2.
  - For the wheat mildew example: $\hat{p} = \hat{p}_2 - \hat{p}_1 + 0.5 = 0.7 - 0.3 + 0.5 = 0.9$
- There are several nonparametric methods for statistically comparing two groups, but most do not generalize to multiple groups, or factorials, or are not appropriate for ordinal data
  - The approach of this workshop covers all of these situations
  - Relative treatment effects and their differences (e.g., p_1-p_2, p_3-p_4, …) are applicable for all factorials

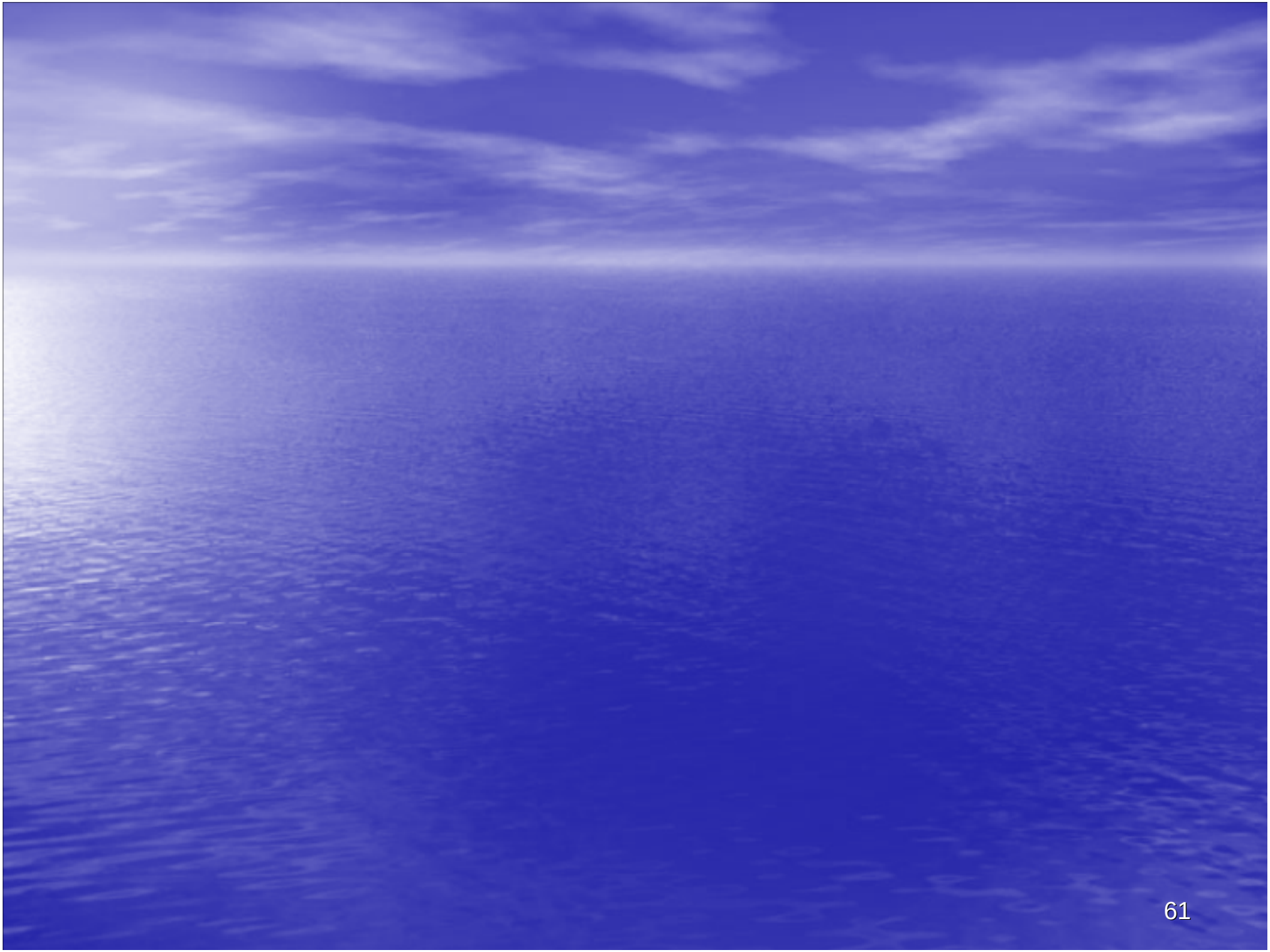
Relative treatment effects

Wheat powdery mildew example:

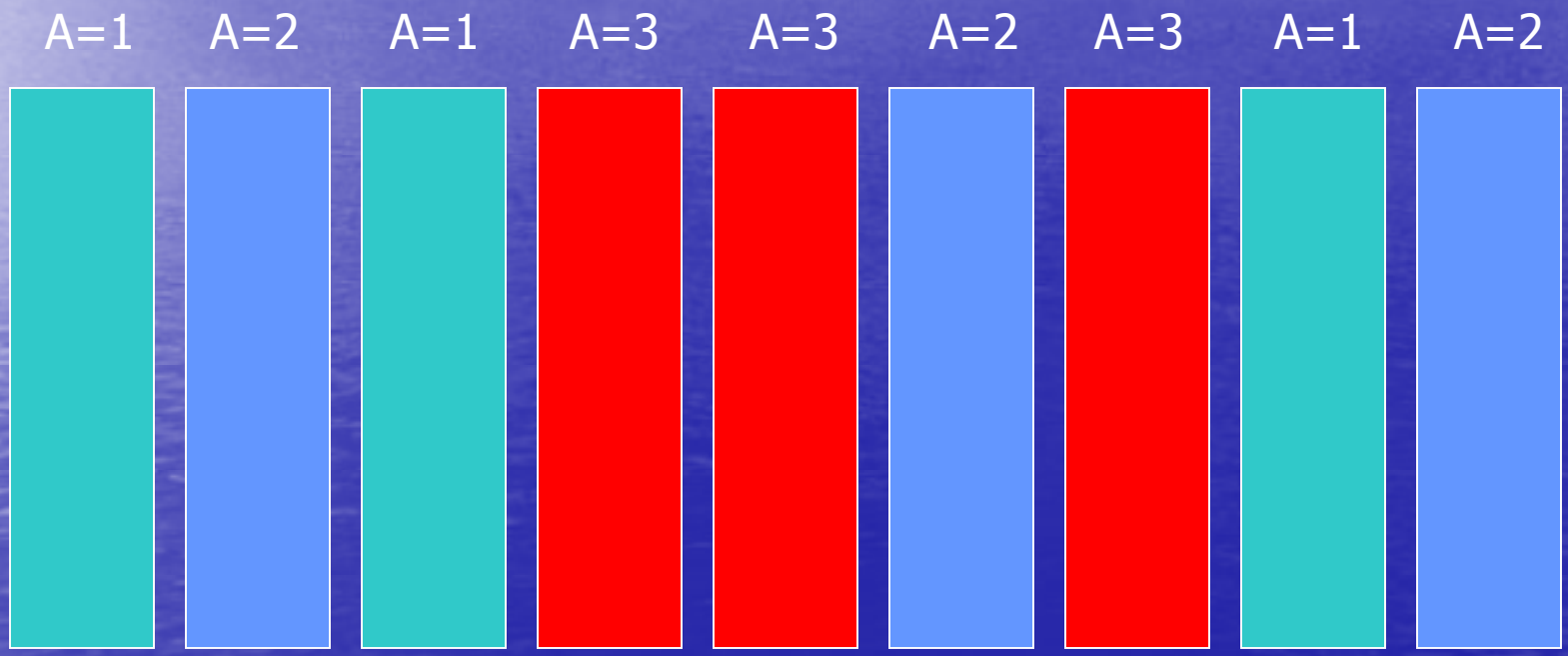
$\overline{R}_{i\bullet}$ $\hat{p}_i$

Cultivar	Median	Mean Rank	Est. Rel. Trt. Eff.
1	8	61.9	0.77
2	5.5	45.2	0.56
3	5.0	34.1	0.42
4	2.0	20.8	0.25





One way layout, completely randomized (Factor A: a=3 treatments; 3 replications)



Nonparametric hypothesis

$$H_0 = F_1 = F_2 = F_3$$

SAS examples

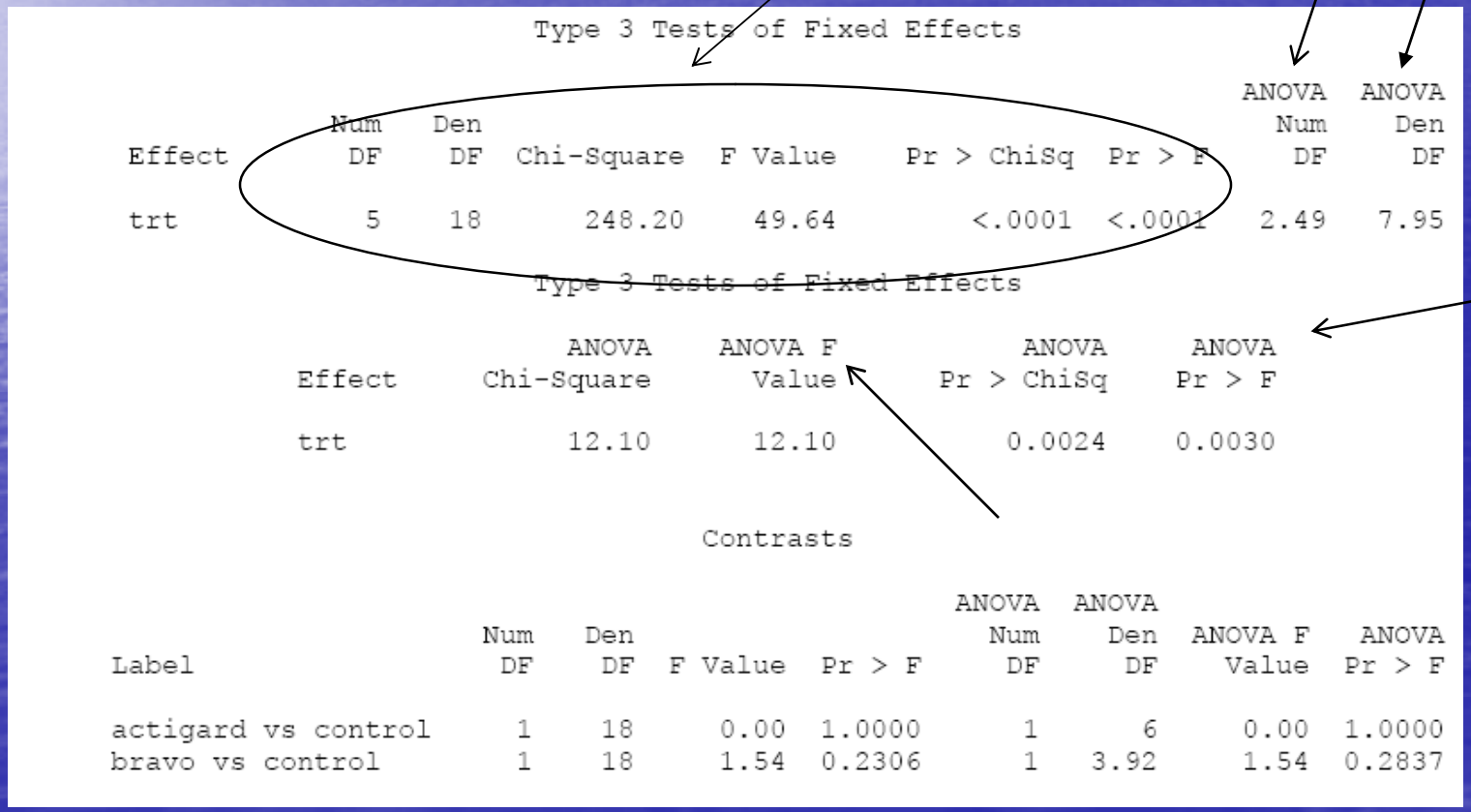
Go to SAS….

Test Statistics

- **"Wald Type Statistic" (WTS)**
  - Asymptotically, has an exact chi-square distribution under the null hypothesis
    - Obtain with the /CHISQ option on the model statement of MIXED
  - But, very large sample sizes are required
  - Do not, in general, use for most data sets
- **"ANOVA Type Statistic" (ATS)**
  - Asymptotically, has an approximate *F* distribution under the null hypothesis
    - Obtain with the ANOVAF option on the procedure statement of MIXED
  - Simulations have shown that this test works (i.e., the statistic has the correct properties) even for very small sample sizes
  - Use for most data sets

One way layout: SAS output

- Version 8.2 output



One way layout: SAS output

- Ver. 9.1 output

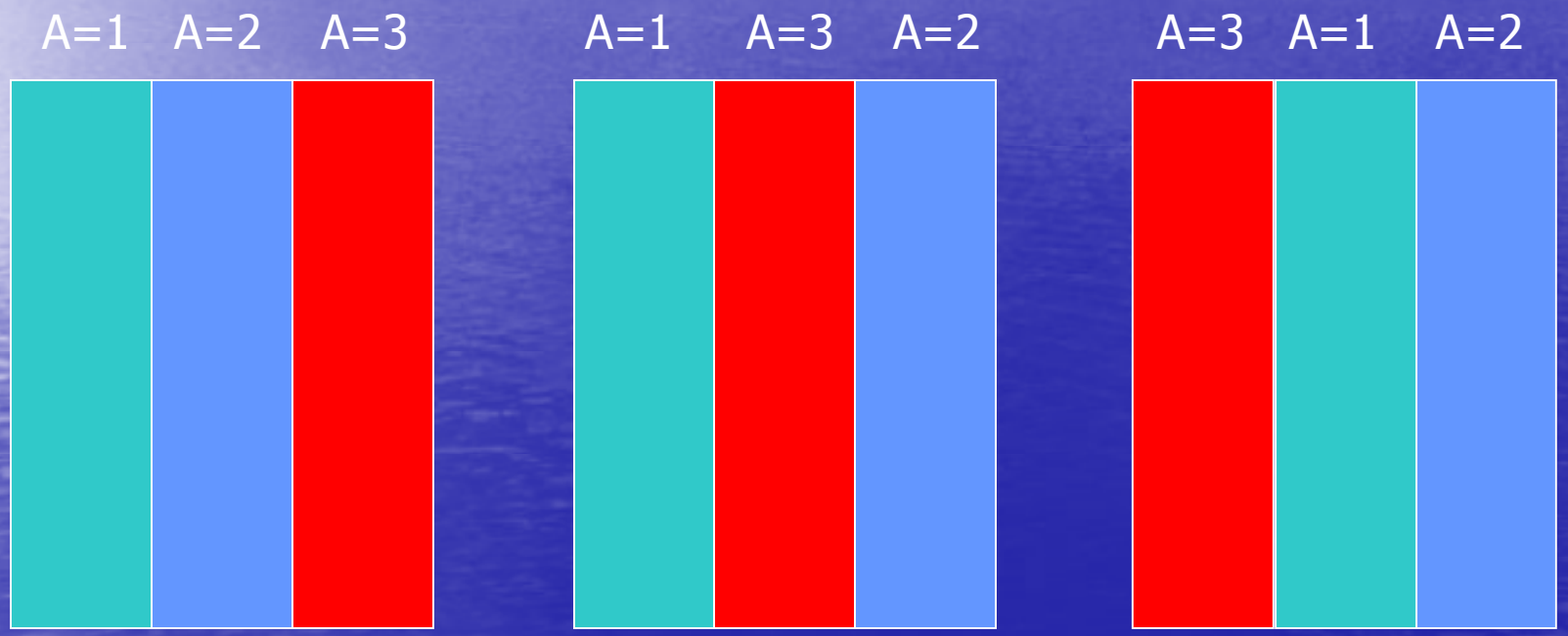
Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
trt	5	18	248.20	49.64	<.0001	<.0001

Type 3 Tests of Fixed Effects

		ANOVA F			
Effect	Num DF	Den DF	Value	Pr > F(DDF)	Pr > F(infty)
trt	2.49	7.95	12.10	0.0030	<.0001

One way layout, with blocking



Dealing with blocking

- Approaches for dealing with blocking are being developed .. Still an active area of current research
- Easiest approach would be to add a random block; statement
- Not accounting for block effects could lead to inflated standard errors

Two way factorial: hypotheses

Parametric hypotheses

$$H_0^{\mu}(A) = \mu_{1.} = \mu_{2.} = ... = \mu_{a.}$$

$$H_0^{\mu}(B) = \mu_{.1} = \mu_{.2} = ...\mu_{.b}$$

$$H_0^{\mu}(AB) = \mu_{ij} + \mu_{..} = \mu_{i.} + \mu_{.j}$$

Generalization

Nonparametric hypotheses

$$H_0^{F}(A): \overline{F}_{1.} = \overline{F}_{2.} = ... = \overline{F}_{a.}$$

$$H_0^{F}(B): \overline{F}_{.1} = \overline{F}_{.2} = ...\overline{F}_{.b}$$

$$H_0^{F}(AB) = F_{ij} + \overline{F}_{..} = \overline{F}_{i.} + \overline{F}_{.j}$$

SAS examples

Go to SAS….

Two way layout: SAS output (vinca)

- Ver. 8.2 output

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F	ANOVA Num DF	ANOVA Den DF
trt	5	72	209.89	41.98	<.0001	<.0001	3.32	36.7
fert	2	72	4.09	2.05	0.1293	0.1367	1.98	36.7
trt*fert	10	72	9.27	0.93	0.5070	0.5142	6.1	36.7

Type 3 Tests of Fixed Effects

Effect	ANOVA Chi-Square	ANOVA F Value	ANOVA Pr > ChiSq	ANOVA Pr > F
trt	18.22	18.22	0.0004	<.0001
fert	2.18	2.18	0.1395	0.1275
trt*fert	1.25	1.25	0.9741	0.3020

Two way layout: SAS output (vinca)

- Ver. 9.1 output

Type 3 Tests of Fixed Effects

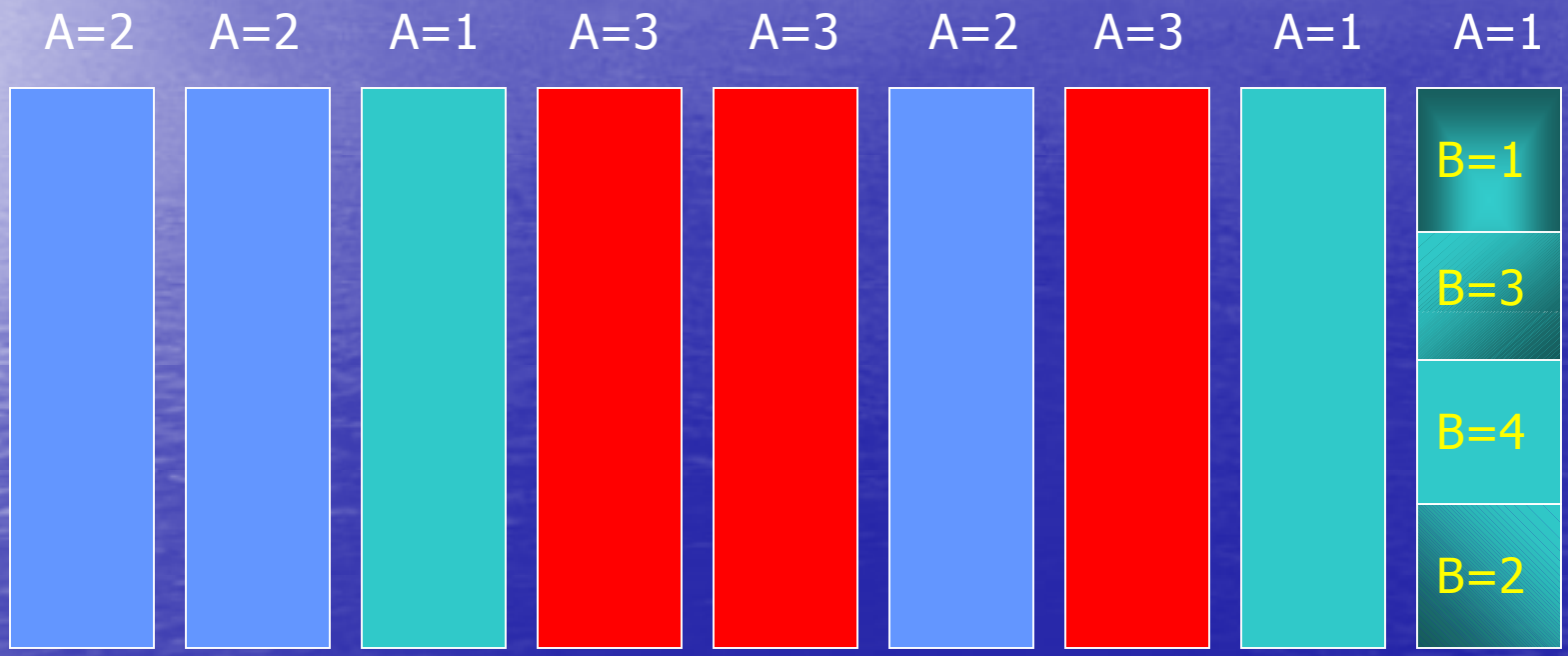
Effect	Num DF	Den DF	Chi-Square	F Value	Pr > ChiSq	Pr > F
trt	5	72	209.89	41.98	<.0001	<.0001
fert	2	72	4.09	2.05	0.1293	0.1367
trt*fert	10	72	9.27	0.93	0.5070	0.5142

Type 3 Tests of Fixed Effects

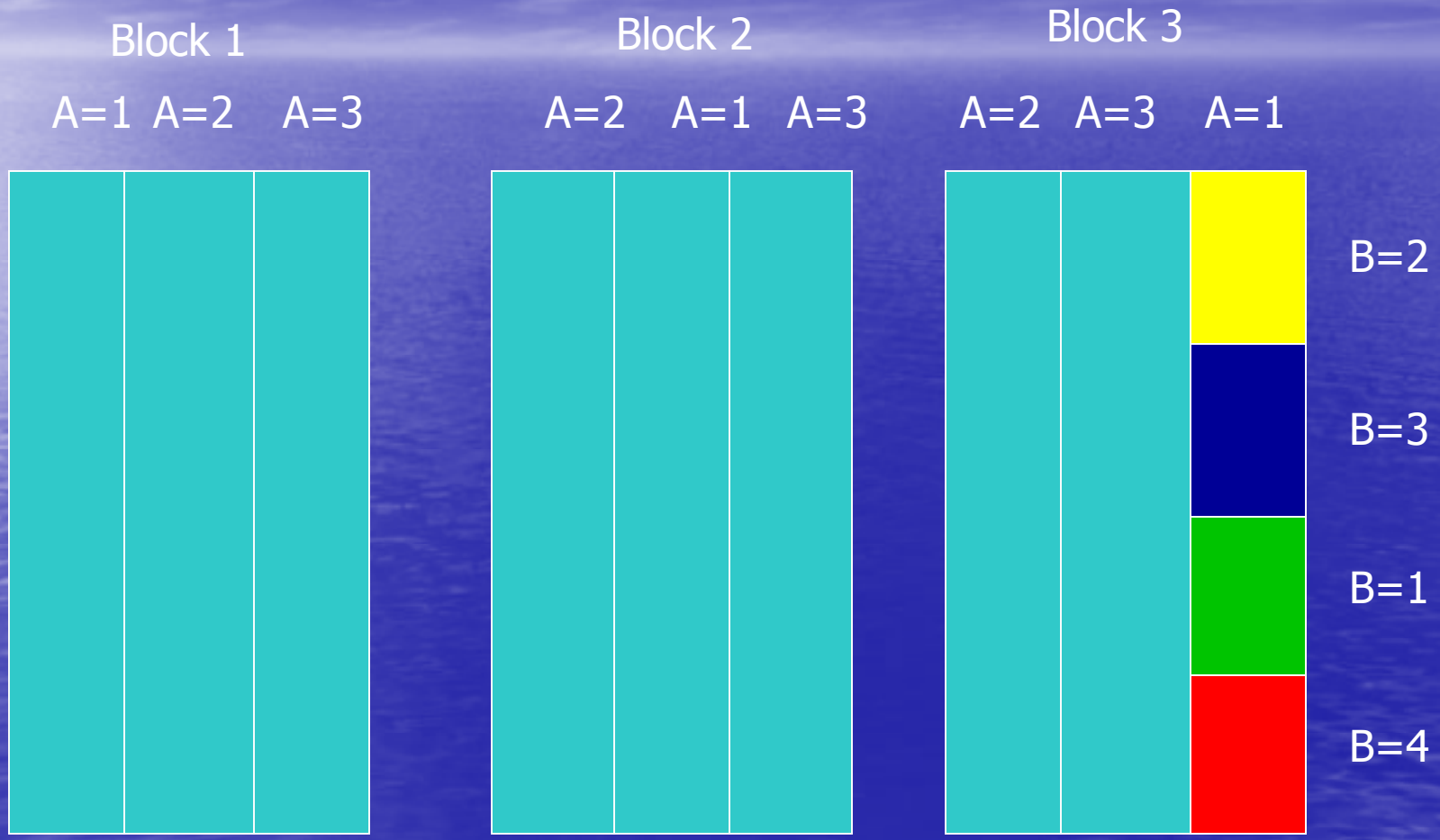
	ANOVA F				
Effect	Num DF	Den DF	Value	Pr > F(DDF)	Pr > F(infty)
trt	3.32	36.7	18.22	<.0001	<.0001
fert	1.98	36.7	2.18	0.1275	0.1131
trt*fert	6.1	36.7	1.25	0.3020	0.2745

Split plot layout

Factor A: a=3 treatments [whole plot];
Factor B: b=4 treatments [sub-plot];
3 replications)



Split plot layout, with blocking



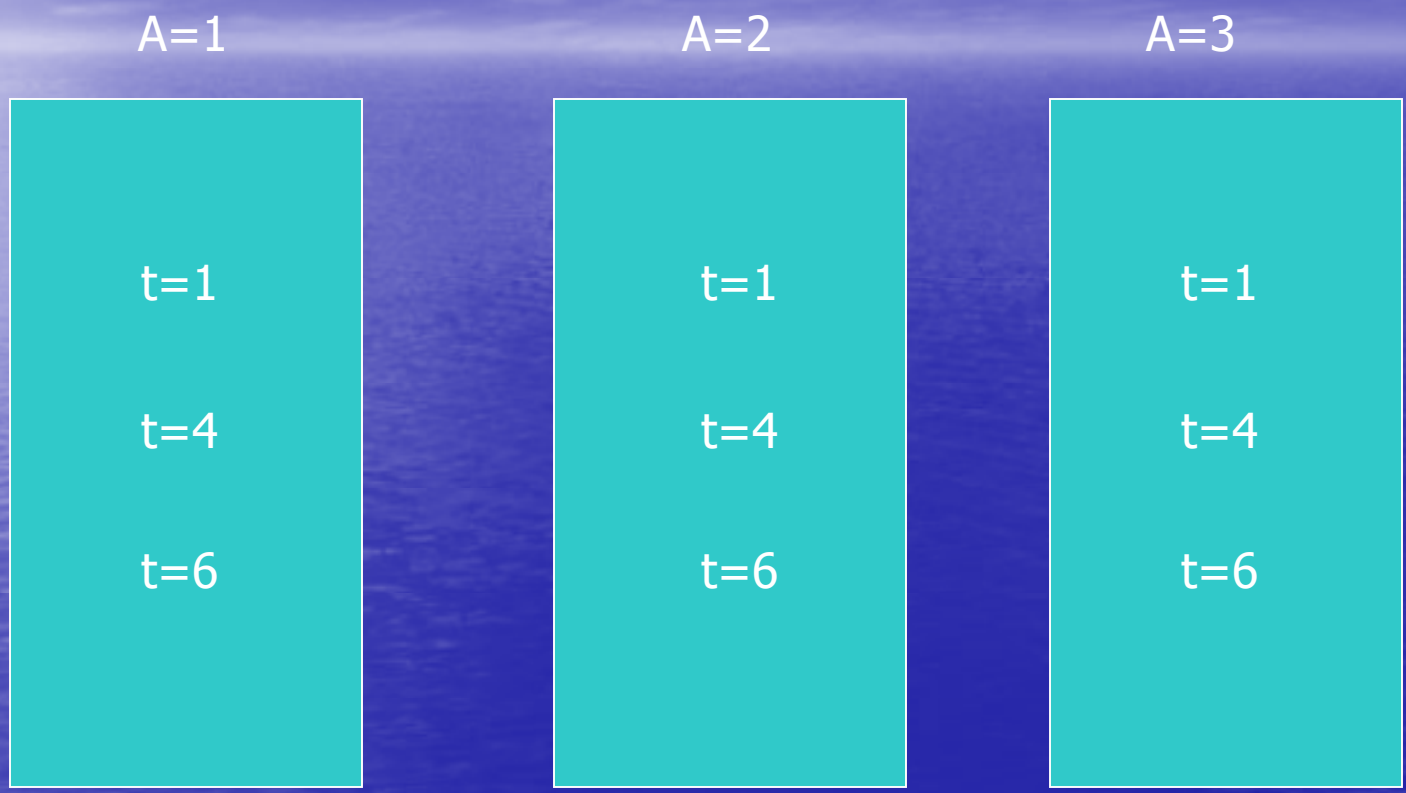
SAS examples

Go to SAS….

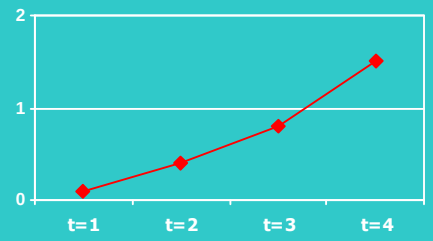
Significance level corresponding to ATS (ANOVA Type Statistic)

- **For crossed factors (1-way, 2-way, etc.)**
  - Use *calculated* numerator and denominator degrees of freedom (Num DF and Den DF)
- **For split plots and repeated measures**
  - Use *calculated* numerator degrees of freedom (Num DF) and *infinite* denominator degrees of freedom ("infty")
    - However, an improved significance level *can* be obtained for the whole-plot (the independent groups) by using *calculated* denominator degrees of freedom (Den DF)
      - Caution: for small sample sizes, one may need to run PROC MIXED a second time to obtain the correct Den DF for whole plot – see comments in e-Xtra.

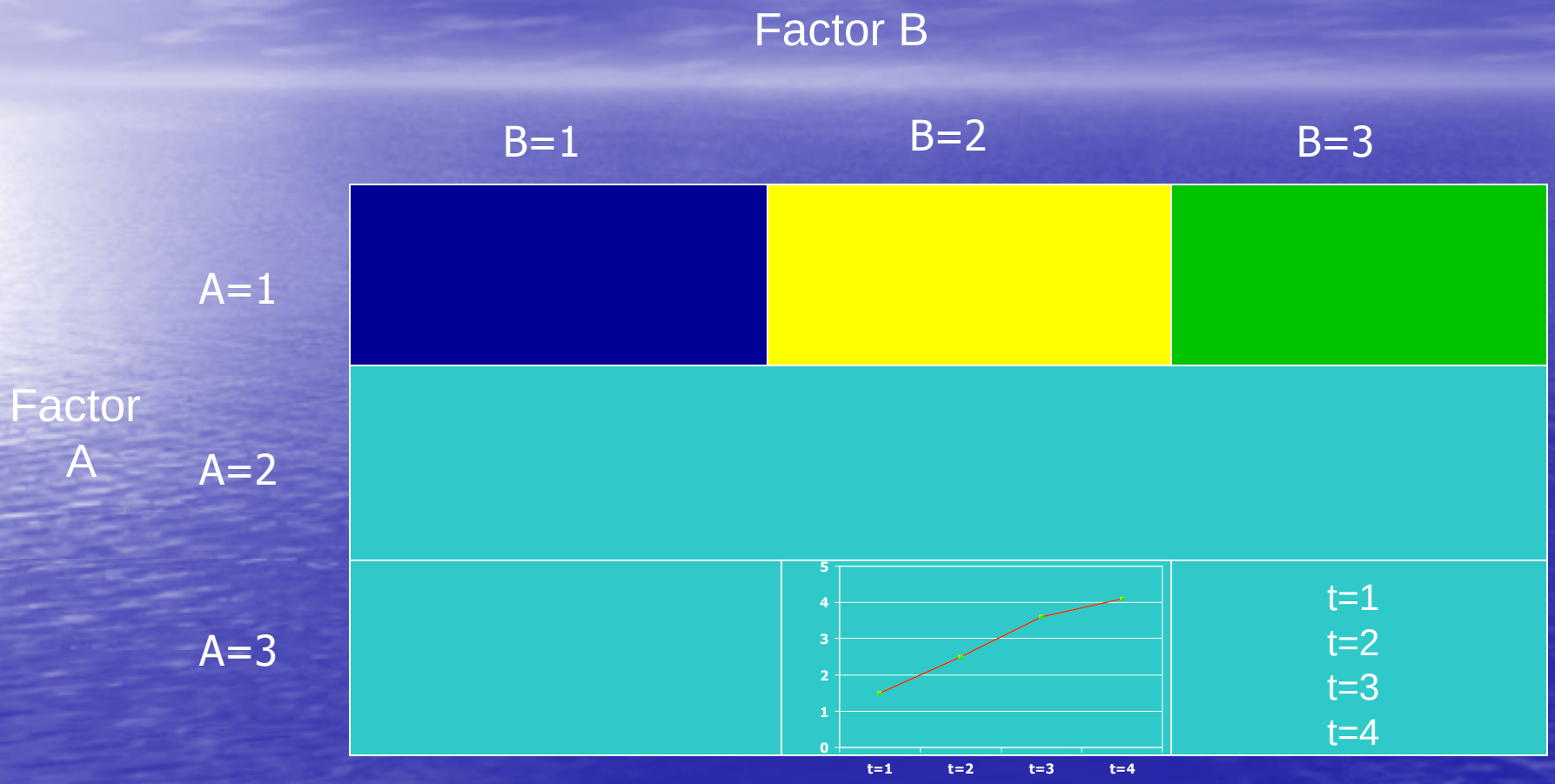
One way repeated measures



Two-way factorial repeated measures

Factor B	Factor A: A=1	Factor A: A=2
B=1		t=1 t=2 t=3 t=4
B=2	t=1 t=2 t=3 t=4	t=1 t=2 t=3 t=4

Split plot repeated measures



Articles which have used the Brunner nonparametric approach to ordinal data:

- Zhao et al. 2004. Plant Dis. 88:1033-1039
- Khan et al. 2004. Plant Dis. 88:280-286
- Dillard et al. 2005. Plant Dis. 89:700-704

Presenting your results…

Table 1. Median, mean rank ($\overline{R}_{ij\bullet}$), and relative treatment effects ($\hat{p}_{ij}$) along with 95% confidence intervals (CI) for snap bean pod russet severity ratings in relation to bean variety and isolate of *Plectosporium tabacinum*

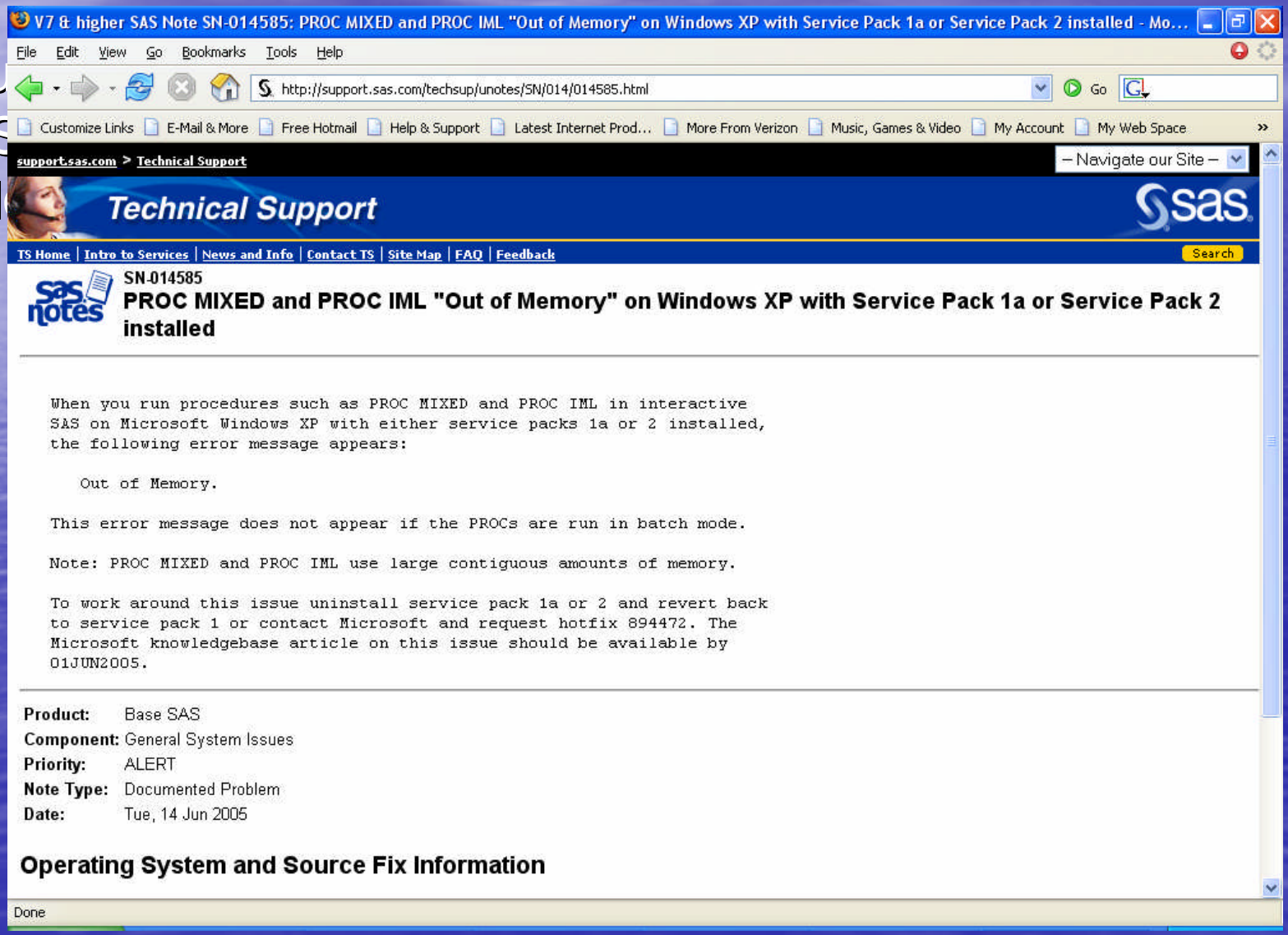
Variety	Isolate[a]	Median[b]	$\overline{R}_{ij\bullet}$	$\hat{p}_{ij}$	95% CI for $\hat{p}_{ij}$
Brio	Control	0.0	17.8	0.161	(0.122, 0.215)
	985	4.5	77.7	0.721	(0.631, 0.794)
	991	1.0	35.6	0.328	(0.270, 0.393)
	1038	2.0	70.5	0.654	(0.441, 0.814)
	1040	5.0	81.6	0.758	(0.679, 0.820)
Gold Mine	Control	0.0	13.0	0.117	(0.103, 0.133)
	985	3.0	59.3	0.549	(0.411, 0.679)
	991	0.0	17.8	0.161	(0.099, 0.259)
	1038	2.0	55.3	0.512	(0.436, 0.588)
	1040	3.0	59.4	0.551	(0.481, 0.618)
Hercules	Control	0.0	13.0	0.117	(0.103, 0.133)
	985	5.0	81.8	0.759	(0.654, 0.836)
	991	1.0	32.0	0.294	(0.272, 0.318)
	1038	5.0	78.8	0.732	(0.543, 0.854)
	1040	5.0	80.3	0.746	(0.578, 0.855)

[a] Controls were sprayed with sterile distilled water. Isolates 985, 1038, and 1040 were obtained from snap bean pods. Isolate 991 was from zucchini.

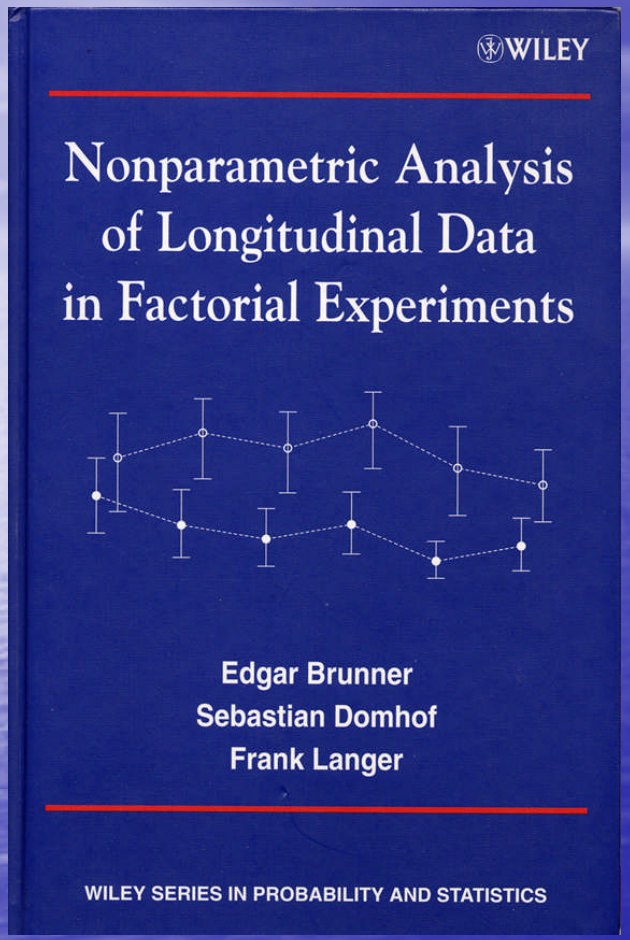
[b] Severity of russet on pods was assessed visually on an ordinal 0 to 9 scale, where 0 = no symptoms and 9 = 100% of the pod surface covered with russet.

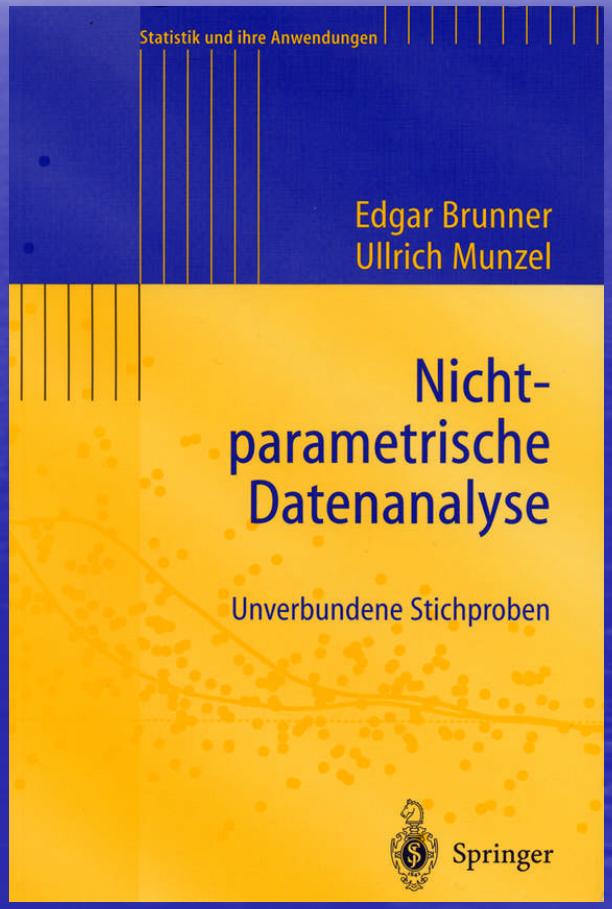
A possible caveat…

- If you
  and S
  probl



Reference books





Some acknowledgements

- Data sets
  - Ann Cobb, NYSAES
  - Camilla Yandoc, USDA ARS USHRL
  - Bob Harveson, Julie Schimelfenig, U. of Nebraska
  - Reza Darvish, INP-ENSAT France
  - Jeff Stein, SDSU
  - Pat Lipps, Jim Chatfield, Dan Herms, OSU

- Slides
  - Beth Gugino, NYSAES
  - Julie Miranda, NCSU
  - Peter Rogers, University of Wisconsin

Future workshops: stay tuned!

- Bayesian analysis (2006)
  - A. Mila & J. Yuen
- Repeated measures analysis
- Spatial statistics