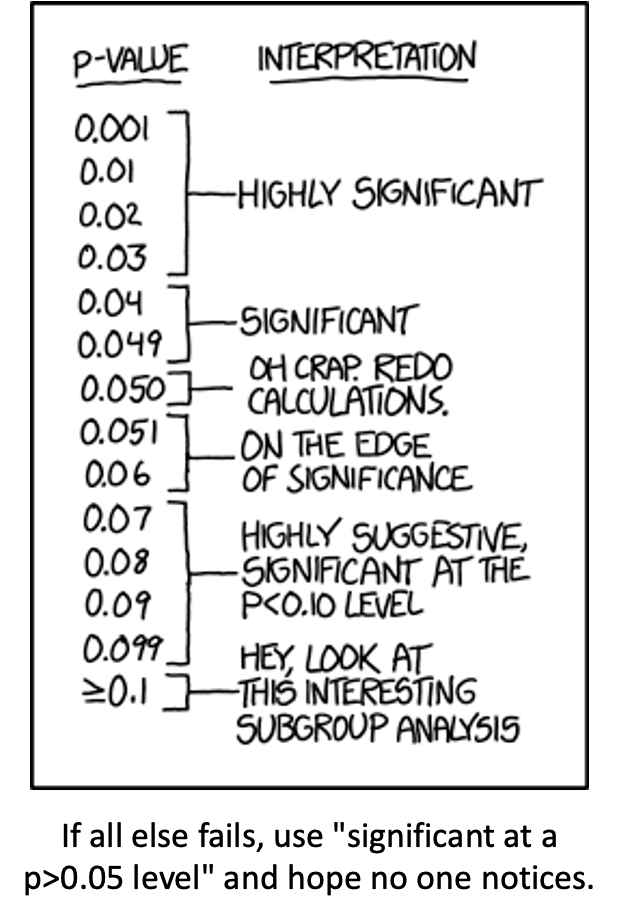
Lecture 13 - Multifactor ANOVA

Bill Perry

# Lecture 13: Review

Multifactor ANOVA

* Example
* Linear model
* Analysis of variance
* Null hypotheses
* Interactions and main effects
* Unequal sample size
* Assumptions



# Lecture 13: 2 Factor or 2 Way ANOVA

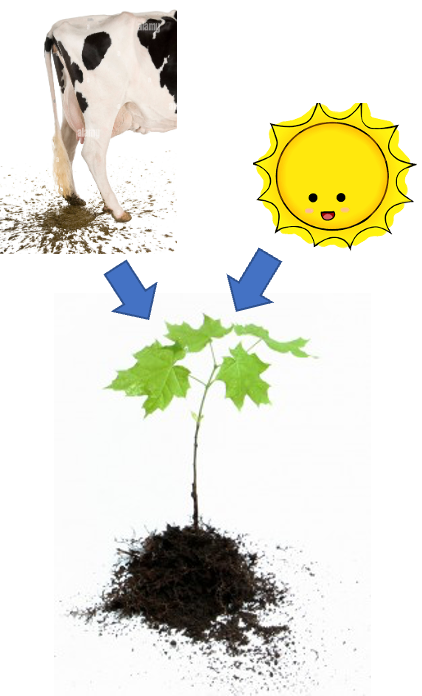
Often consider more than 1 factor (independent categorical variable):

* reduce unexplained variance
* look at interactions

2-factor designs (2-way ANOVA) very common in ecology

* Can have more factors (e.g., 3-way ANOVA)
  + interpretation tricky…

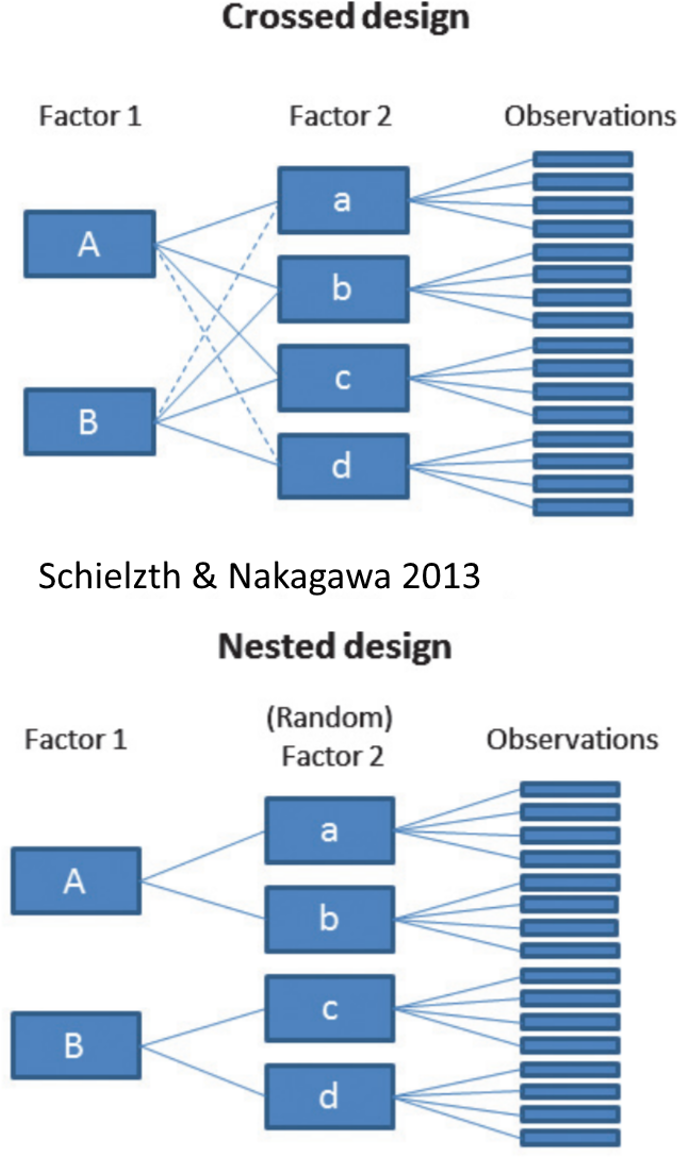
Most multifactor designs: nested or factorial



# Factorial Versus Nested Designs

Consider two factors: A and B

* Factorial/crossed: every level of B in every level of A
* Nested/hierarchical: levels of B occur only in 1 level of A



# Lecture 13: Nested ANOVA Overview

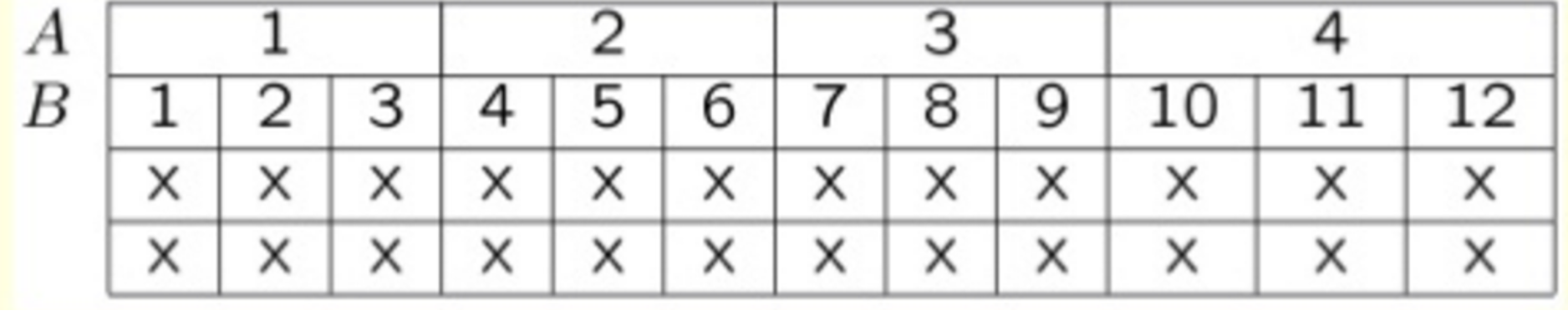
Nested design examples

* Nested designs
* Linear model
* Analysis of variance
* Null hypotheses
* Unbalanced designs
* Assumptions

# Nested Designs Overview

Nested Designs:

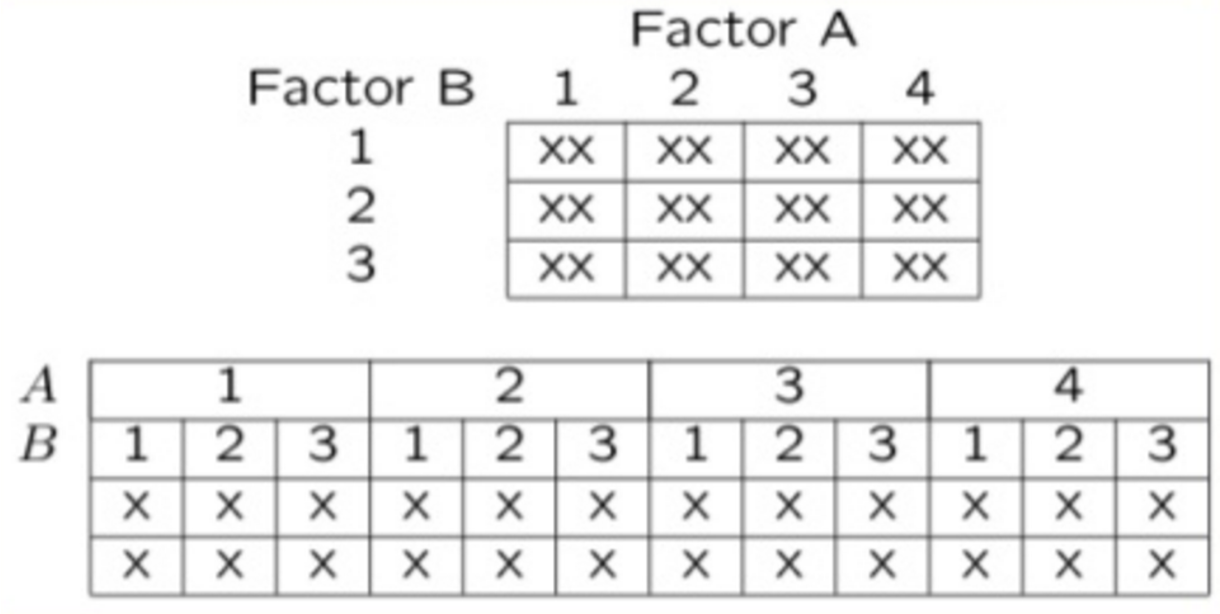
* Factor A usually fixed
* Factor B usually random



# Factorial Designs Overview

Factorial Designs:

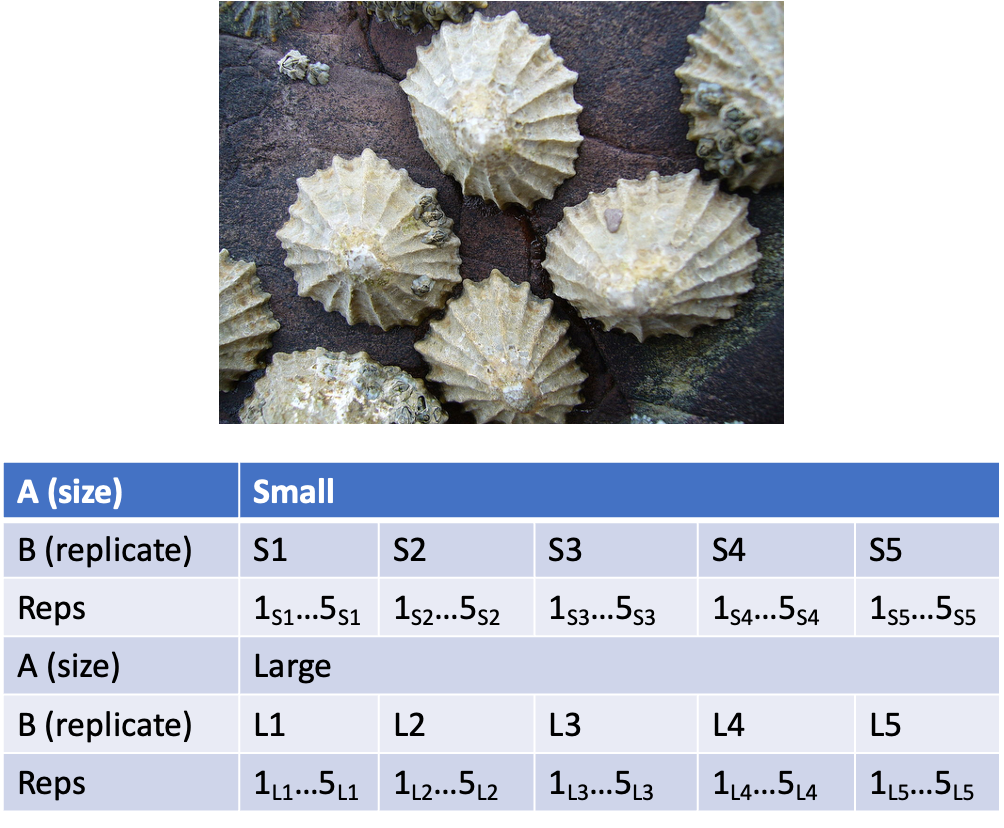
* Both factors typically fixed (but not always)



# Nested Design Example: Limpet Growth

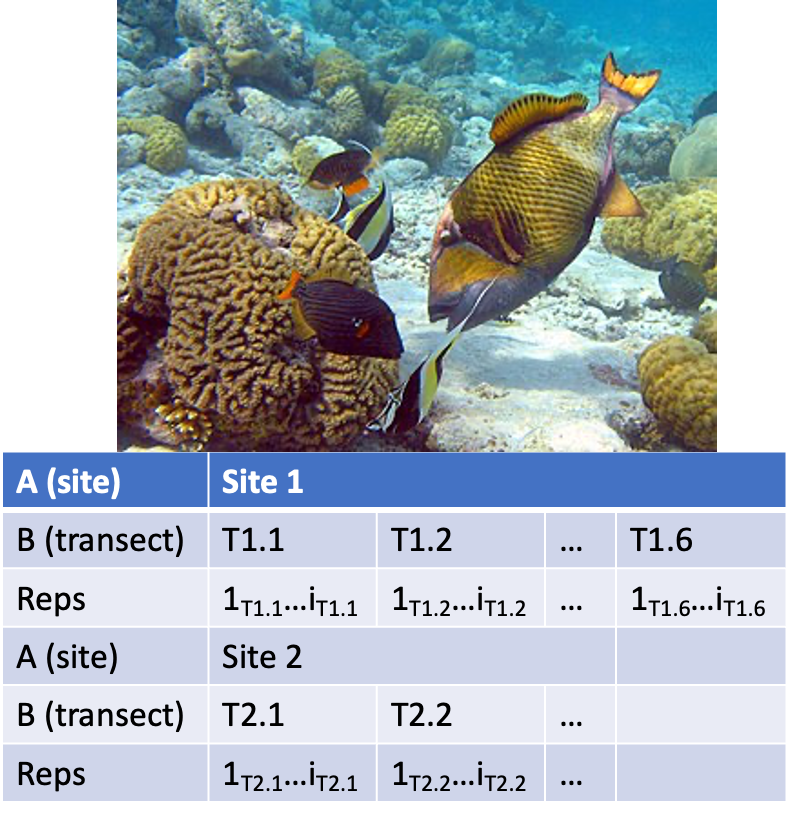
Study on effects of enclosure size on limpet growth:

* 2 enclosure sizes (factor A)
* 5 replicate enclosures (factor B)
* 5 replicate limpets per enclosure



# Nested Design Example: Reef Fish

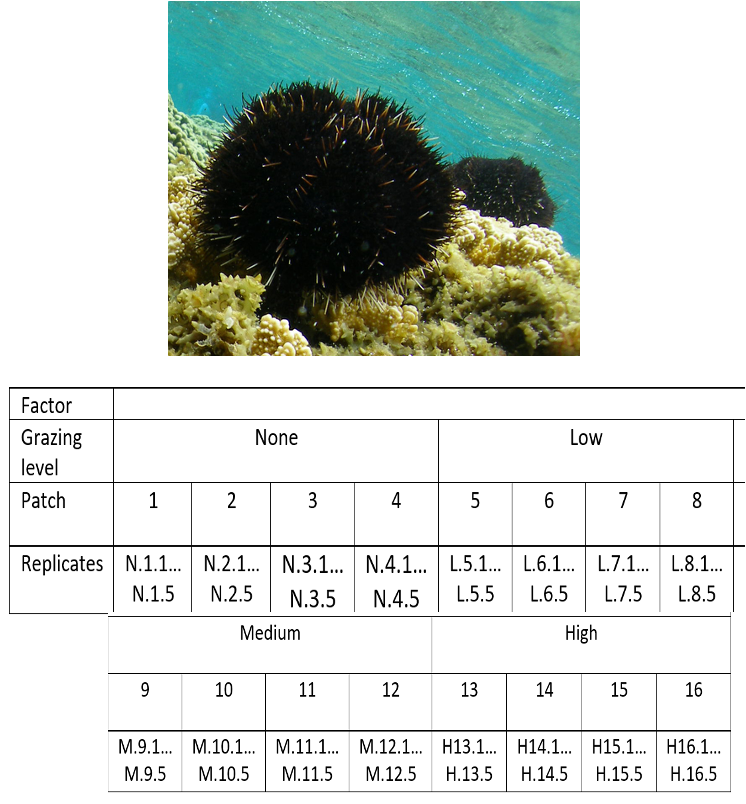
Study on reef fish recruitment: 5 sites (factor A) 6 transects at each site (factor B) replicate observations along each transect



# Nested Design Example: Sea Urchin Grazing

Effects of sea urchin grazing on biomass of filamentous algae:

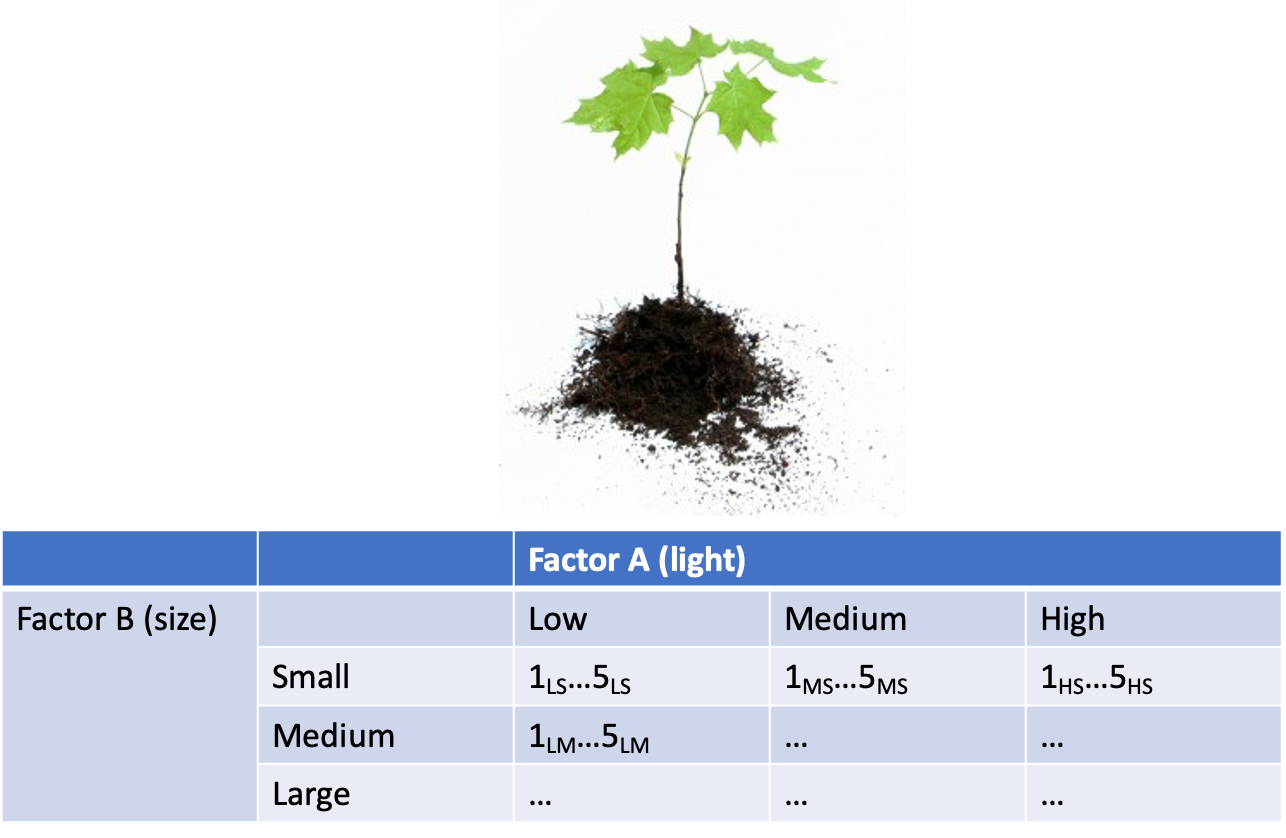
* 4 levels of urchin grazing: none, L, M, H
* 4 patches of rocky bottom (3-4 m2) nested in each level of grazing
* 5 replicate quadrats per patch



# Factorial Design Example: Seedling Growth

Effects of light level on growth of seedlings of different size:

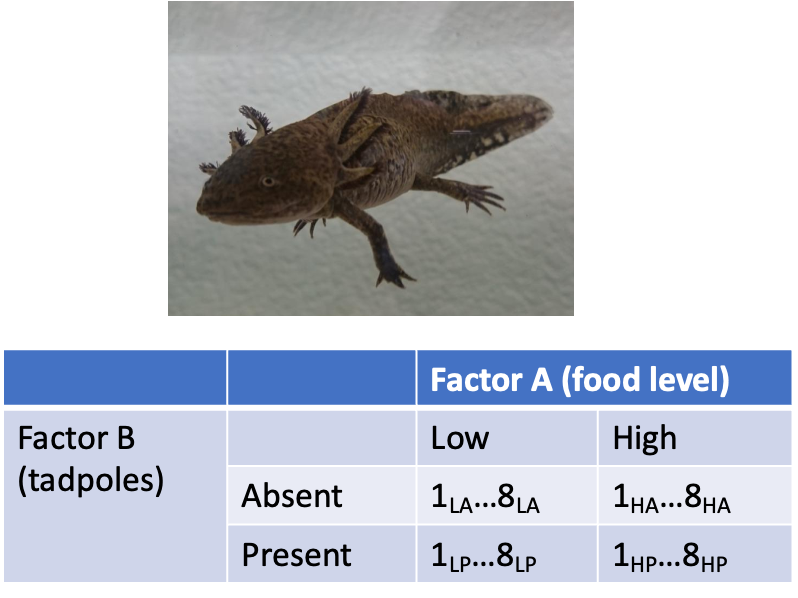
* 3 light levels (factor A)
* 3 size classes (factor B)
* 5 replicate seeding in each cell



# Factorial Design Example: Salamander Growth

Effects of food level and tadpole presence on larval salamander growth

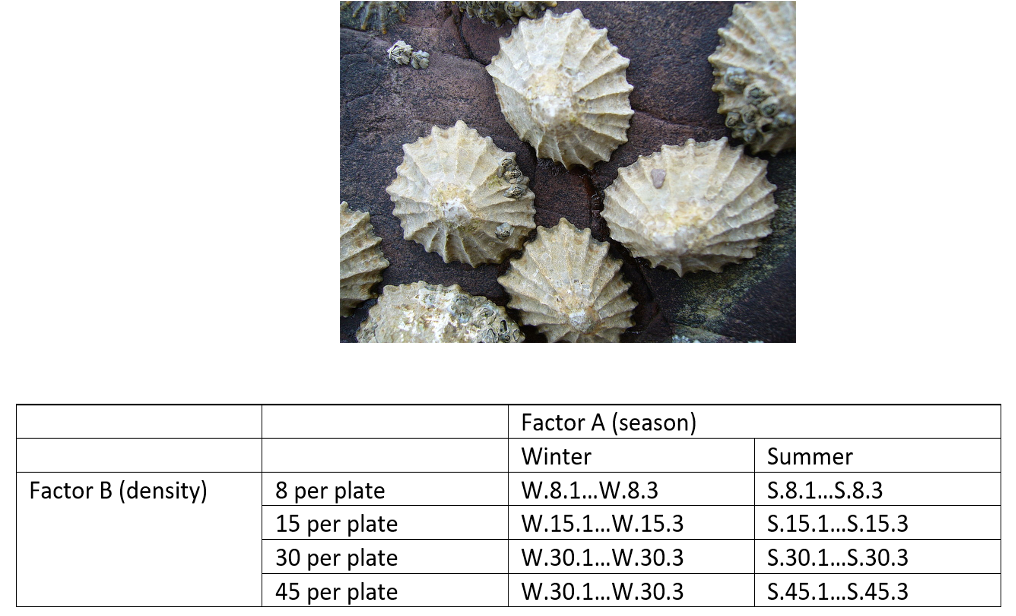
* 2 food levels (factor A)
* presence/absence of tadpoles (factor B)
* 8 replicates in each cell



# Factorial Design Example: Limpet Fecundity

Effect of season and density on limpet fecundity.

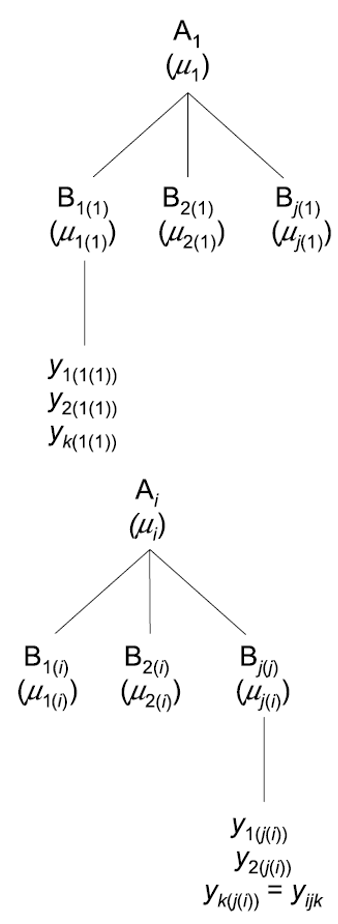
* 2 seasons (factor A)
* 4 density treatments (factor B)
* 3 replicates in each cell



# Nested Design: Linear Model Structure

Consider a nested design with:

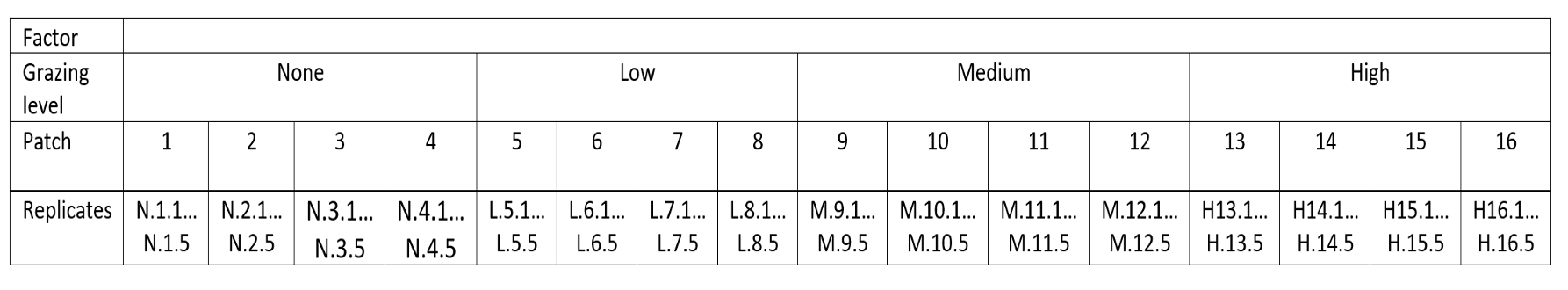
* p levels of factor A (i= 1…p) (e.g., 4 grazing levels)
* q levels of factor B (j= 1…q), nested within each level of A (e.g., 4 - diff. patches per grazing level)
* n replicates (k= 1…n) in each combination of A and B (5 replicate - quadrats in each patch in each grazing level)



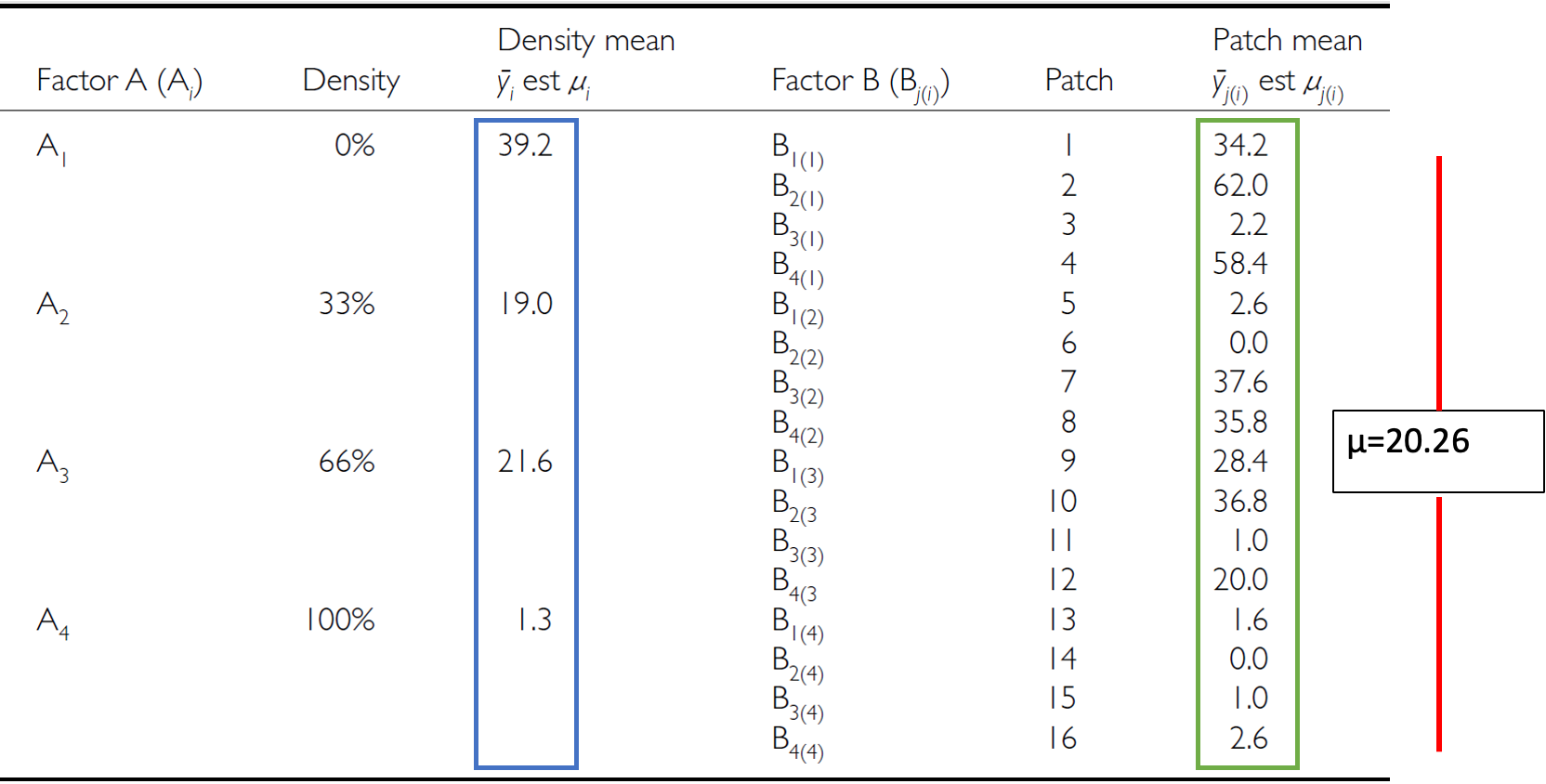
# Calculating Means in Nested Design

Can calculate several means:

* overall mean (across all levels of A and B)= ȳ;
* a mean for each level of A (across all levels of B in that A)= ȳi;
* a mean for each level of B within each A= ȳj(i)



# Nested Design Means Visualization



# Nested Design Linear Model

The linear model for a nested design is:

Where:

* is the response variable
  + value of the k-th replicate in j-th level of B in the i-th level of A
  + (algal biomass in 3rd quadrat, in 2nd patch in low grazing treatment)
* is the overall mean
  + (overall average algal biomass)

# Fixed Effects in Nested Model

The linear model for a nested design is:

* is the fixed effect of factor
* (difference between average biomass in all low grazing level quadrats and overall mean)
* is the random effect of factor nested within factor
* usually random variable, measuring variance among all possible levels of B within each level of A
* (variance among all possible patches that may have been used in the low grazing treatment)

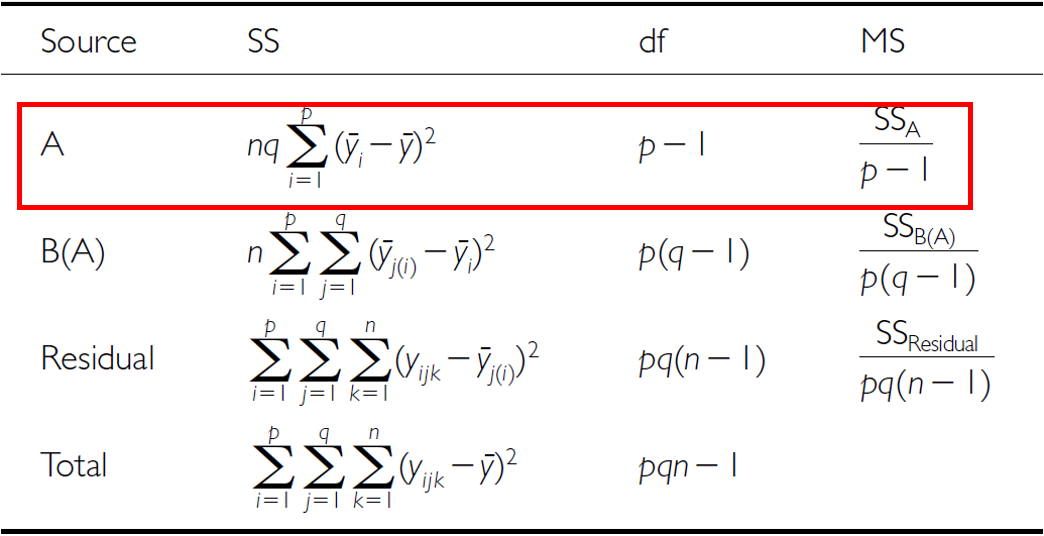
# Error Term in Nested Model

The linear model for a nested design is:

* is the error term
* αi: is the effect of the ith level of A: µi- µ
* unexplained variance associated with the kth replicate in jth level of B in the ith level of A
* (difference bw observed algal biomass in 3rd quadrat in 2nd patch in low grazing treatment and predicted biomass - average biomass in 2nd patch in low grazing treatment)

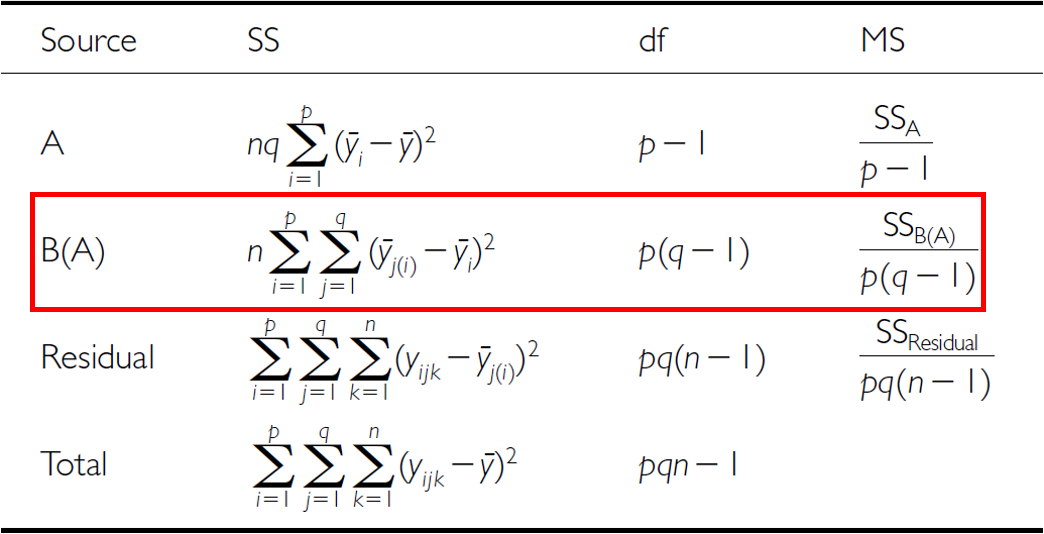
# Analysis of Variance: SSA

As before, partition the variance in the response variable using SS SSA is SS of differences between means in each level of A and overall mean



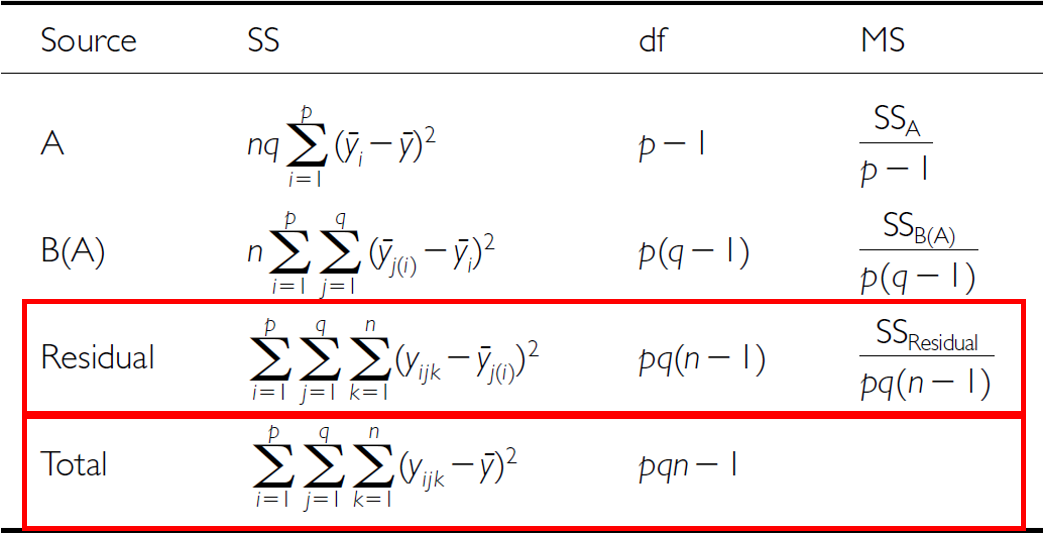
# Analysis of Variance: SSB

SSB is SS of difference between means in each level of B and the mean of corresponding level of A summed across levels of A

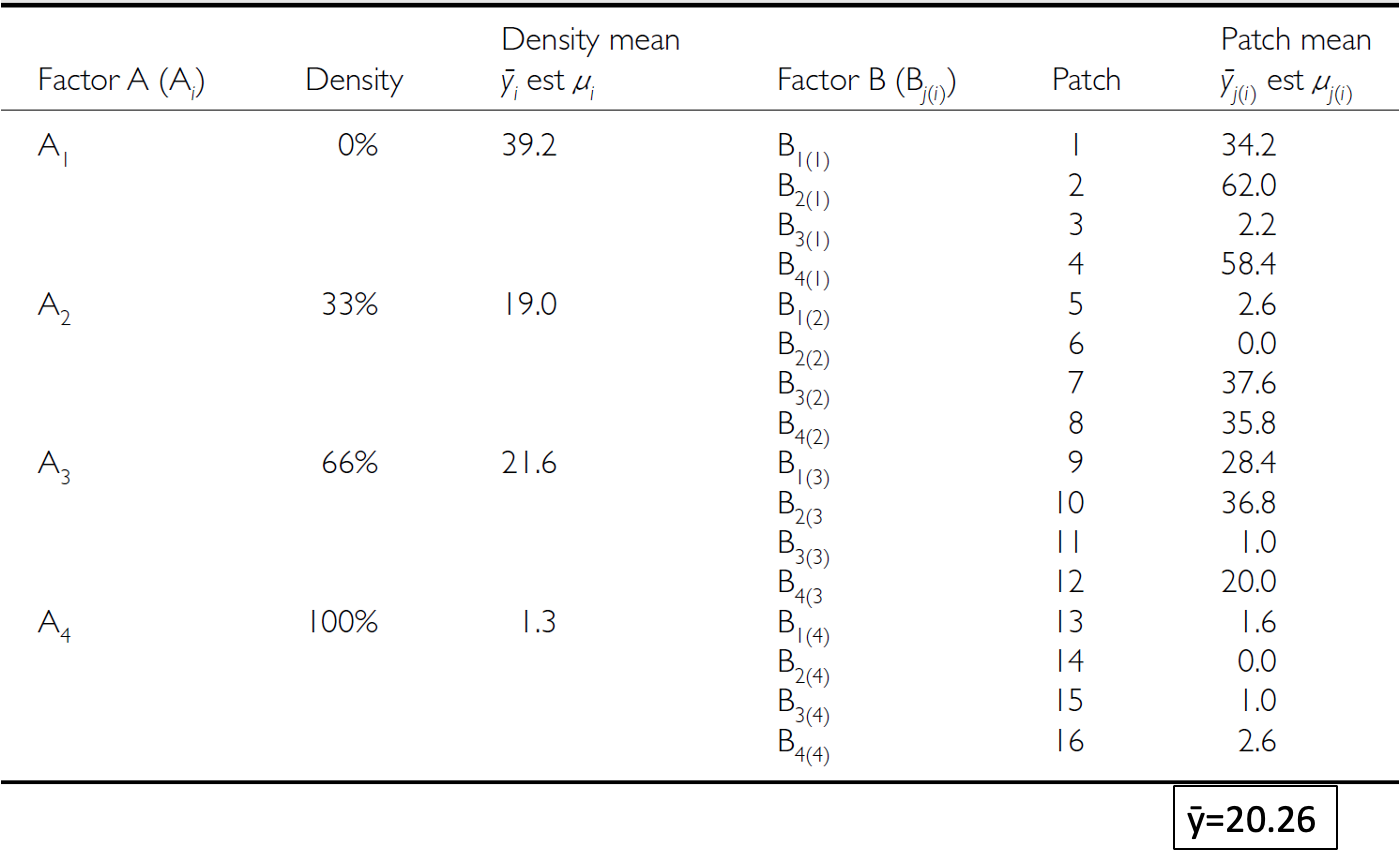


# Analysis of Variance: Residual and Total

* SSresid is difference bw each observation and mean for its level of factor B, summed over all observations
* SStotal = SSA + SSB + SSresid
* SS can be turned into MS by dividing by appropriate df



# Analysis of Variance Table

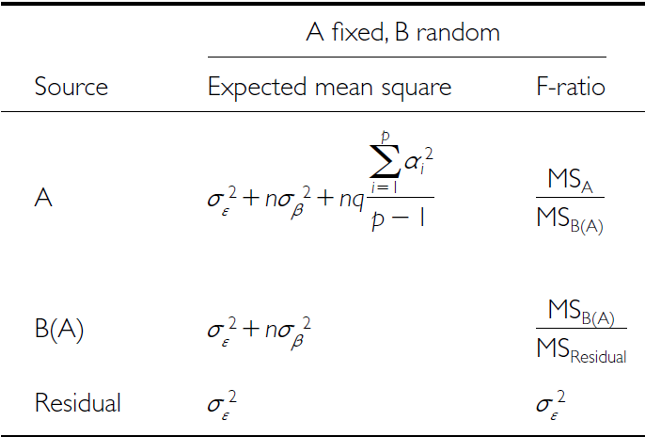


# Null Hypotheses: Factor A

Two hypotheses tested on values of MS:

1. no effects of factor A

* Assuming A is fixed:
* Ho(A): µ1= µ2= µ3=…. µi= µ
* Same as in 1-factor ANOVA, using means from B factors nested within each - level of A
* (no difference in algal biomass across all levels of grazing: µnone= - µlow= µmed= µhigh)

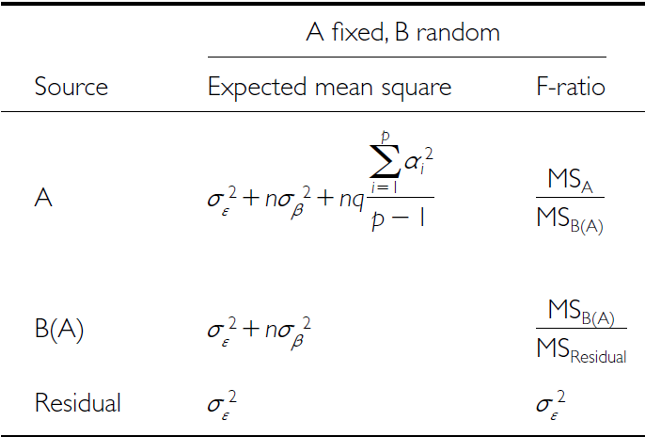


# Null Hypotheses: Factor B

Two hypotheses tested on values of MS:

1. No effects of factor B nested in A

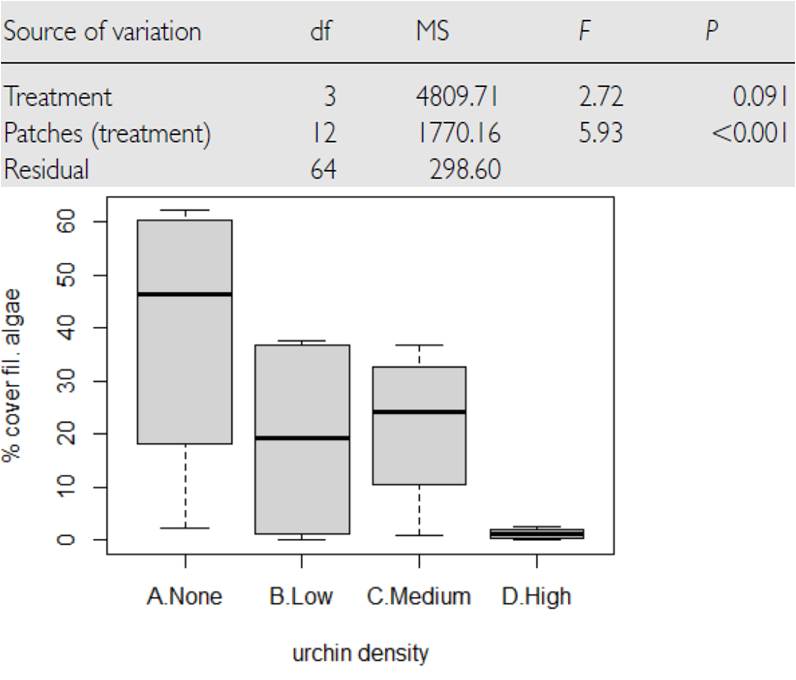
* Assuming B is random:
* Ho(B): σβ2= 0 (no variance added due to differences between all possible - levels of B)
* (no variance added due to differences between patches)



# Conclusions from Analysis

**Conclusions?**

“significant variation between replicate patches within each treatment, but no significant difference in amount of filamentous algae between treatments”

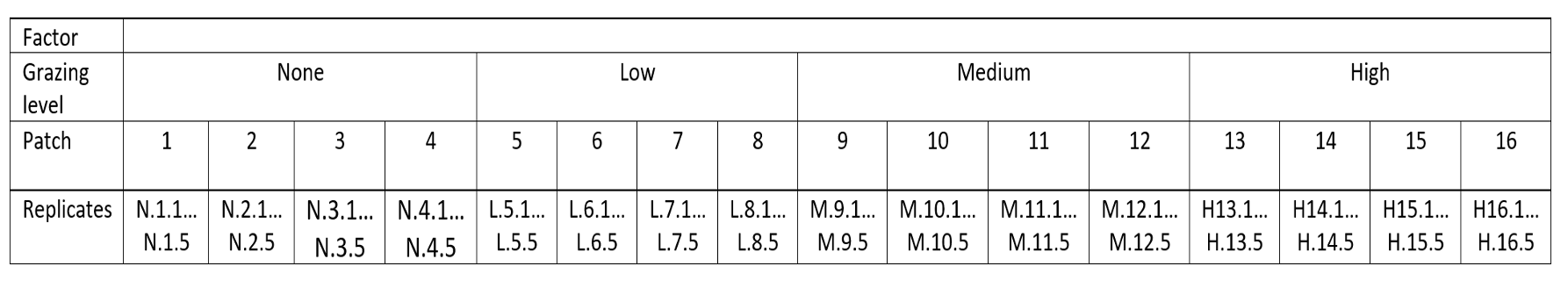


# Unbalanced Nested Designs

Unequal sample sizes can be because of:

* uneven number of B levels within each A
* uneven number of replicates within each level of B

Not a problem, unless have unequal variance or large deviation from - normality



# Nested Design Assumptions

As usual, we assume

* equal variance
* normality
* independence of observations

Equal variance + normality need to be assessed at both levels:

* Since means for each level of B within each A are used for the H-test about A, need to assess whether those means meet normality and equal variance
* Examine residuals for H-test about B
* Transformations can be used