Module 3C: ANOVA & Correlation

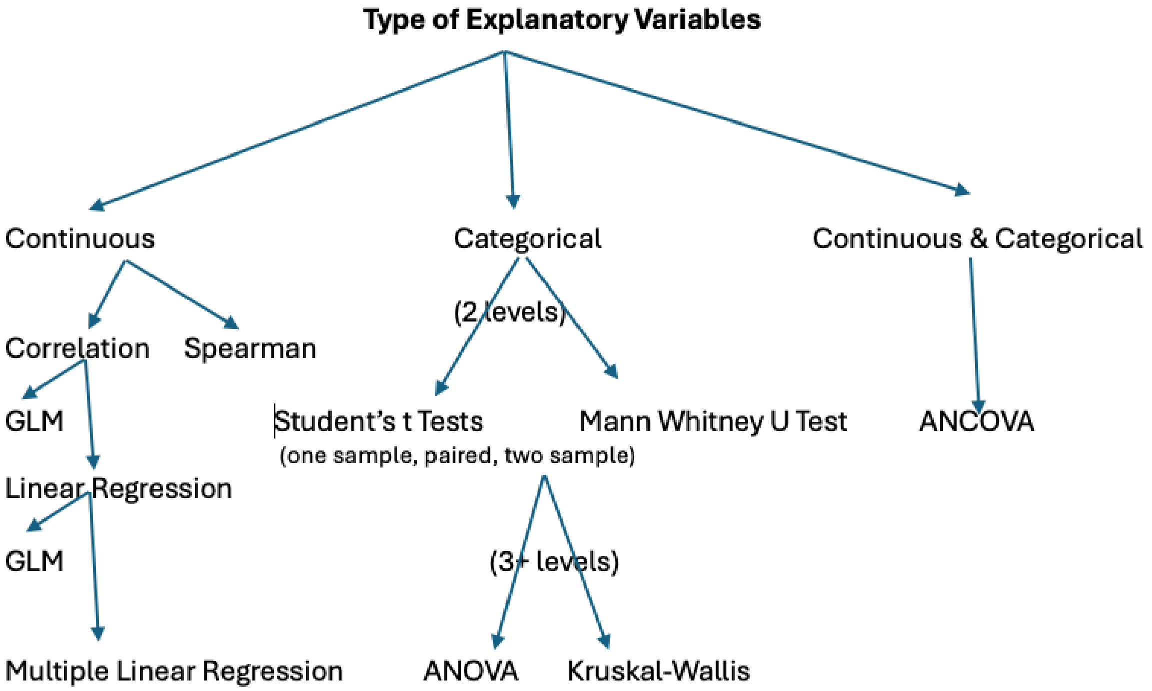
Assigning signal and noise to variation

Agenda:

1. ANOVA: Nuts & Bolts

2. Worked Example
   A. **One way ANOVA**
   B. Post-hoc tests: Tukey-Kramer
   C. Kruskal-Wallis (nonparametric)

3. Linear Correlation
   A. Spearman's rank



Kruskal-Wallis Test:

o A non-parametric test similar to a single factor ANOVA

o Uses the **ranks** of the data points; tests **medians** not means

- Data points are **not** compared, their ranks are!

*Using **ranks** is what frees us from having to assume normality since all distributions have similar predictions about ranks*

- All group samples are random samples
- Distribution of the variable has the same shape in every population
- Small samples lead to little power but when n is large, Kruskal-Wallis has the same power as ANOVA

o ***H***, sampling distribution is χ^2 with df = k - 1

Researchers want to know whether **activity level** differs across three mouse strains (B6, BALB/c, CAST) after 12 weeks on the same diet. Activity is scored on a 1–10 scale.

Here are the activity scores:

Strain A: B6 (n = 4)

M1: 3.0
M2: 3.5
M3: 6.0
M4: 6.5

Strain B: BALB/c (n = 4)

M5: 4.0
M6: 4.5
M7: 7.0
M8: 7.5

Strain C : CAST (n = 4)

M9: 5.0
M10: 5.5
M11: 8.0
M12: 8.5

Researchers want to know whether **activity level** differs across three mouse strains (B6, BALB/c, CAST) after 12 weeks on the same diet. Activity is scored on a 1–10 scale.

Activity is not normally distributed, and you need to use a non-parametric test:

Strain A – B6 (n = 4)	Strain B – BALB/c (n = 4)	Strain C – CAST (n = 4)
M1: 3.0	M5: 4.0	M9: 5.0
M2: 3.5	M6: 4.5	M10: 5.5
M3: 6.0	M7: 7.0	M11: 8.0
M4: 6.5	M8: 7.5	M12: 8.5

Step 1: Formulate the null/alternate hypothesis

Ho: The median activity level among the three mouse strains is equal (the independent samples all have the same central tendency and therefore come from the same underlying population)

Ha: The median activity level among the three mouse strains is not equal (at least one of the independent samples does not have the same central tendency and therefore originates from a different population)

Researchers want to know whether **activity level** differs across three mouse strains (B6, BALB/c, CAST) after 12 weeks on the same diet. Activity is scored on a 1–10 scale.

Activity is not normally distributed, and you need to use a non-parametric test:

Strain A – B6 (n = 4)	Strain B – BALB/c (n = 4)	Strain C – CAST (n = 4)
M1: 3.0	M5: 4.0	M9: 5.0
M2: 3.5	M6: 4.5	M10: 5.5
M3: 6.0	M7: 7.0	M11: 8.0
M4: 6.5	M8: 7.5	M12: 8.5

Step 1: Formulate the null/alternate hypothesis

Ho: The median activity level among the three mouse strains is equal

Ha: The median activity level among the three mouse strains is not equal

Step 2: Assumptions and test

Kruskal- Wallis test; observations are independent, Strains are ordinal with continuous dependent variables, and the strains come from populations with similar distribution shape.

Researchers want to know whether **activity level** differs across three mouse strains (B6, BALB/c, CAST) after 12 weeks on the same diet. Activity is scored on a 1–10 scale. Activity is not normally distributed, and you need to use a non-parametric test:

Strain A – B6 (n = 4)	**Strain B – BALB/c (n = 4)**	**Strain C – CAST (n = 4)**
M1: 3.0	M5: 4.0	M9: 5.0
M2: 3.5	M6: 4.5	M10: 5.5
M3: 6.0	M7: 7.0	M11: 8.0
M4: 6.5	M8: 7.5	M12: 8.5

Step 1: Formulate the null/alternate hypothesis

Ho: The median activity level among the three mouse strains is equal

Ha: The median activity level among the three mouse strains is not equal

Step 2: Assumptions and test

Kruskal- Wallis test; observations are independent, Strains are ordinal with continuous dependent variables, and the strains come from populations with similar distribution shape.

3.0, 3.5, 4.0,4.5, 5.0, 5.5, 6.0, 6.5, 7.0 7.5, 8.0, 8.5

R_A = 1+2+7+8 = 18

R_B = 3+4+9+10= 26

R_C = 5+6+11+12=34

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$

Researchers want to know whether **activity level** differs across three mouse strains (B6, BALB/c, CAST) after 12 weeks on the same diet. Activity is scored on a 1–10 scale. Activity is not normally distributed, and you need to use a non-parametric test:

Strain A – B6 (n = 4)	Strain B – BALB/c (n = 4)	Strain C – CAST (n = 4)
M1: 3.0	M5: 4.0	M9: 5.0
M2: 3.5	M6: 4.5	M10: 5.5
M3: 6.0	M7: 7.0	M11: 8.0
M4: 6.5	M8: 7.5	M12: 8.5

Step 1: Formulate the null/alternate hypothesis

Ho: The median activity level among the three mouse strains is equal

Ha: The median activity level among the three mouse strains is not equal

Step 2: Assumptions and test

Kruskal- Wallis test; 3.0, 3.5, 4.0,4.5, 5.0, 5.5, 6.0, 6.5, 7.0 7.5, 8.0, 8.5

N=12

k=3 groups

$\overline{R_A}$=18/4 = 4.5

$$\overline{R_B} = \frac{26}{4} = 6.5$$

$$\overline{R_C} = \frac{34}{4} = 8.5$$

$$E_R = \frac{N+1}{2} = 13/2 = 6.5$$

$$\sigma^2 = \frac{N^2-1}{12} = 143/12$$

R_A = 1+2+7+8 = 18

R_B = 3+4+9+10= 26

R_C = 5+6+11+12=34

$$H = \frac{N-1}{N}\sum_{i=1}^{k}\frac{n_i(\bar{R}-E_R)^2}{\sigma^2}$$

$$= \frac{11}{12}\sum_{i=1}^{k}\frac{4(\frac{18}{4}-(12+1)/2)^2}{(12^2-1)/12} + \frac{4(\frac{26}{4}-(12+1)/2)^2}{(12^2-1)/12} + \frac{4(\frac{34}{4}-(12+1)/2)^2}{(12^2-1)/12} = 2.61$$

Researchers want to know whether **activity level** differs across three mouse strains (B6, BALB/c, CAST) after 12 weeks on the same diet. Activity is scored on a 1–10 scale. Activity is not normally distributed, and you need to use a non-parametric test:

Strain A – B6 (n = 4)	**Strain B – BALB/c (n = 4)**	**Strain C – CAST (n = 4)**
M1: 3.0	M5: 4.0	M9: 5.0
M2: 3.5	M6: 4.5	M10: 5.5
M3: 6.0	M7: 7.0	M11: 8.0
M4: 6.5	M8: 7.5	M12: 8.5

Step 1: Formulate the null/alternate hypothesis

Ho: The median activity level among the three mouse strains is equal

Ha: The median activity level among the three mouse strains is not equal

Step 2: Assumptions and test

Kruskal- Wallis test; 3.0, 3.5, 4.0,4.5, 5.0, 5.5, 6.0, 6.5, 7.0 7.5, 8.0, 8.5

H = 2.61

Step 3: Critical value

There are **df**=k-1=**2** and you use a χ^2 table to find the cut-off value.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	$\boxed{5.991}$	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Researchers want to know whether **activity level** differs across three mouse strains (B6, BALB/c, CAST) after 12 weeks on the same diet. Activity is scored on a 1–10 scale. Activity is not normally distributed, and you need to use a non-parametric test:

Strain A – B6 (n = 4)	**Strain B – BALB/c (n = 4)**	**Strain C – CAST (n = 4)**
M1: 3.0	M5: 4.0	M9: 5.0
M2: 3.5	M6: 4.5	M10: 5.5
M3: 6.0	M7: 7.0	M11: 8.0
M4: 6.5	M8: 7.5	M12: 8.5

Step 1: Formulate the null/alternate hypothesis

Ho: The median activity level among the three mouse strains is equal

Ha: The median activity level among the three mouse strains is not equal

Step 2: Assumptions and test

Kruskal- Wallis test; 3.0, 3.5, 4.0,4.5, 5.0, 5.5, 6.0, 6.5, 7.0 7.5, 8.0, 8.5

H = 2.61

Step 3: Critical value

There are **df**=k-1=**2** and you use a χ^2 table to find the cut-off value which is **5.991** for α=0.05.

Step 4: Decision

Fail to reject the null hypothesis

Fixed Effects: The groups *are* the question

- Also called Model 1 ANOVA
  - What we have been using so far
- Different categories of explanatory variable are predetermined and repeatable
  - **Results cannot be generalizable**
  - Example: <u>specific</u> drug treatment, <u>specific</u> diets, <u>specific</u> season

Random Effects: The groups are a source of noise in the system you are modeling

- Also called Model 2 ANOVA
- Different categories of explanatory variable are *randomly sampled from a larger population of groups*
  - **Results are generalizable;** conclusions reached about difference among groups can be generalized to the whole population
  - Example: family in a study about resemblance of IQ
    - Chose a <u>random family</u> in a population of families
    - Family: group
    - Replicates: different children within each family
  - **The population and not the particular families involved is the target of study**

Fixed Effects: **The groups *are* the question**

- Also called Model 1 ANOVA
  - What we have been using so far
- Different categories of explanatory variable are predetermined and repeatable
  - **Results cannot be generalizable**
  - Example: specific drug treatment, specific diets, specific season

Do you want to estimate the effect of this specific group (fixed), or treat the group as one random example drawn from a broader population (random)?

Example: You compare **Chow vs. HFD vs. Low-Fat** diets.

- You care about *those* diets, not some random sample of all possible diets in the universe.
- You want to estimate: μ_{Chow}, μ_{LFD}, μ_{HFD}
- These diet means *are the whole point of the study* → **fixed effects.**

Random Effects: The groups are a source of noise in the system you are modeling

- Also called Model 2 ANOVA
- Different categories of explanatory variable are *randomly sampled from a larger population of groups*
  - **Results are generalizable;** conclusions reached about difference among groups can be generalized to the whole population
  - **The population and not the particular families involved is the target of study**

Do you want to estimate the effect of this specific group (fixed), or treat the group as one random example drawn from a broader population (random)?

Example: Your mice come from **several litters** or **multiple cages**, but you don't care about comparing litter #1 vs litter #2.

- You treat **litter** or **cage** as a random effect because:
- Litters are not scientifically interesting
- They create real variability
- You want to **control for** that variability without estimating each litter's mean

Random effects represent "background noise" you want to account for because it changes the variance, not because you want to measure each group

Example: Suppose you measure **PPARG expression** across three strains: B6, BALB/c, CAST

If your goal is:

A. *"How do PPARG levels differ between these specific strains?"*

→ **Strain = Fixed effect**

B. *"Strain is just a nuisance variable; these 3 strains are a random sample of wild genetic diversity."*

→ **Strain = Random effect**

The same factor can be fixed or random *depending on the scientific question*.

Quick heuristic:

Ask: “Do I want to estimate the mean of each group?”

Yes → Fixed effect
No → Random effect

Ask: “Do these group levels represent all the possibilities or just a sample?”

All (or all that matter) →Fixed effect
Just a sample of many possible levels →Random effect