Multivariate Analysis of Variance (MANOVA)

1 Introduction: Why MANOVA?

Many research problems involve more than one response variable. For example, a psychologist might study treatments affecting both anxiety and depression, or a biologist might compare fertilizers by measuring plant height, leaf area, and root weight simultaneously.

If we analyze each outcome separately with ANOVA, we risk:

1. Inflated Type I error (multiple testing).
2. Ignoring correlations among outcomes, which might contain critical information.

MANOVA extends ANOVA by considering a vector of dependent variables simultaneously. It evaluates whether the mean vectors across groups are equal, accounting for correlations among variables.

2 One-Way MANOVA

2.1 Model Setup

- Suppose we have k groups, each with n_i independent observations.
- Each observation is a p-dimensional response vector:

$$\mathbf{y}_{ij} = \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{ij}, \quad \boldsymbol{\varepsilon}_{ij} \sim N_p(\mathbf{0}, \Sigma).$$

- Null hypothesis:

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k.$$

2.2 Hypothesis Matrices

- Within-group (Error) matrix:

$$\mathbf{E} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)(\mathbf{y}_{ij} - \bar{\mathbf{y}}_i)'.$$

- Between-group (Hypothesis) matrix.

$$\mathbf{H} = \sum_{i=1}^{k} n_i (\bar{\mathbf{y}}_i - \bar{\mathbf{y}})(\bar{\mathbf{y}}_i - \bar{\mathbf{y}})'.$$

- These partition the total variation:

$$\mathbf{T} = \mathbf{H} + \mathbf{E}.$$

Where:

- **T (Total SSCP Matrix):** Represents the total variation and covariation of the dependent variables across all observations, regardless of group membership.
- **H (Hypothesis SSCP Matrix):** Represents the variation and covariation between the group means. It is the part of the total variation that can be explained by the differences among the groups.
- **E (Error SSCP Matrix):** Represents the variation and covariation within each group. This is the unexplained, or random, variation.

The goal of MANOVA is to determine if the H matrix is significant relative to the E matrix. If the variation between groups (H) is substantially larger than the variation within groups (E), the null hypothesis of no group differences is rejected.

2.3 Test Statistics

Unlike ANOVA, which uses a single F-statistic, MANOVA uses several multivariate test statistics to evaluate the significance of the H and E matrices. The most common are:

- **Wilks' Lambda (Λ):** This is the most widely reported statistic. It is the ratio of the determinant of the error SSCP matrix (|E|) to the determinant of the total SSCP matrix (|T|). It ranges from 0 to 1, with a smaller value indicating a greater effect of the independent variable on the dependent variables. A value close to 1 suggests no significant difference between the group mean vectors. The statistic is often converted to an F-statistic for hypothesis testing.
- **Pillai's Trace, Hotelling-Lawley Trace, and Roy's Largest Root:** These are alternative test statistics that are also calculated from the eigenvalues of the

H·E−1 matrix. They each have different strengths, but Wilks' Lambda is the most common and robust.

All MANOVA tests derive from the eigenvalues λ_r of $E^{-1}H$:

1. Wilks' Lambda: likelihood ratio test for MANOVA

$$\Lambda = \frac{\det(\mathbf{E})}{\det(\mathbf{E}+\mathbf{H})} = \prod_r \frac{1}{1+\lambda_r}.$$

   Small values of Λ indicate differences in means.

2. Pillai's Trace:

$$V = \sum_r \frac{\lambda_r}{1+\lambda_r}.$$

   Interpreted as proportion of explained variance; robust to violations.

3. Lawley–Hotelling Trace:

$$T = \sum_r \lambda_r.$$

4. Roy's Largest Root:

$$\theta = \max_r \lambda_r.$$

For two groups, all reduce to Hotelling's T^2 test.

Each statistic follows a distribution under H0, often approximated using complex formulas or tables. Roy's test can be least robust when assumptions fail, and Pillai's trace is noted for its robustness to violations (e.g., unequal covariances)

Assumptions & Robustness

To ensure valid results, MANOVA relies on several key assumptions:

- Independence of Observations: The observations within and between groups must be independent. This is typically a design issue, meaning that no individual or case should be in more than one group.
- Multivariate Normality: The dependent variables should follow a multivariate normal distribution within each group. This is a stricter assumption than

univariate normality. Although difficult to check directly, looking at the univariate normality of each dependent variable can provide an indication.

- Homogeneity of Variance-Covariance Matrices: This is the multivariate equivalent of the homogeneity of variances assumption in ANOVA. It assumes that the population variance-covariance matrices for all groups are equal. This is formally tested using Box's M test. A significant Box's M test suggests a violation of this assumption, which can be problematic, especially with unequal sample sizes.

Post-Hoc and Follow-Up Analyses

After detecting a significant MANOVA, follow-up steps typically include:

- Univariate ANOVAs for each dependent variable (with caution).
- Canonical variate analysis: Identify linear combinations (canonical variates) that best separate group means.
- Discriminant analysis: To classify observations and examine separation structure.

Extensions

- MANCOVA (Multivariate ANCOVA): Incorporates covariates to adjust for extraneous variation and improve power. Test statistics are analogous (Wilks', Pillai's, etc.) but include covariate adjustments.
- Repeated Measures / Factorial Designs: Extensions to more complex MANOVA designs

Scenario: Suppose we have three diets, each applied to 10 subjects, measuring two outcomes—weight loss (kg) and blood pressure reduction (mmHg).

- **Groups**: $k = 3$; **Outcomes**: $p = 2$; **Sample size**: $n = 10$.
- **Compute Within and Between Matrices**:
  - $\bar{\mathbf{y}}_i$—mean vector per diet.
  - $\bar{\mathbf{y}}$—overall mean vector.
  - Form $\mathbf{H}$ and $\mathbf{E}$ as defined above.
- **Find eigenvalues** of $\mathbf{E}^{-1}\mathbf{H}$: say $\lambda_1 = 2.5$, $\lambda_2 = 0.8$.
- **Compute Test Stats**:
  - Wilks' Lambda: $\Lambda = \prod(1/(1+\lambda)) = (1/3.5) \times (1/1.8) \approx 0.159$
  - Pillai's trace: $V = 2.5/3.5 + 0.8/1.8 \approx 0.714 + 0.444 = 1.158$
  - LH Trace: $T = 2.5 + 0.8 = 3.3$
  - Roy's root: $\theta = 2.5$

- **Interpretation**:
  - Low Wilks' Lambda suggests strong group differences.
  - Pillai's trace (sum of explained variance proportions) indicates substantial multivariate effect.
  - LH and Roy's also show effect size via eigenvalues.
  - Software then converts these to approximate F-statistics and p-values.
- **Follow-up**:
  - Univariate ANOVA: check which outcomes contribute most.
  - Canonical Discriminant Analysis: find the direction in (weight loss, bp reduction) space best separating diets.
  - Interpret canonical coefficients or loadings.

2.4 Procedure and Interpretation

The general process for conducting and interpreting a MANOVA is as follows:

Check Assumptions: Before running the MANOVA, check for independence, multivariate normality (or at least univariate normality), and homogeneity of variance-covariance matrices using Box's M test.

Conduct the Overall MANOVA Test: Run the analysis and examine the multivariate test statistics (e.g., Wilks' Lambda). If the overall test is statistically significant, it indicates that there is a difference between the mean vectors of at least two groups.

Step-Down Analysis (Follow-up Tests): A significant overall MANOVA result does not tell you *which* dependent variables are responsible for the significant group differences. A common follow-up procedure is a step-down analysis. This involves performing a series of univariate ANOVAs on the dependent variables in a prioritized order.

- The first ANOVA is run on the most important dependent variable.
- The second ANOVA is run on the next dependent variable, but with the first dependent variable acting as a covariate to statistically control for its effect.
- This process continues, with each subsequent ANOVA controlling for all previously entered dependent variables. This helps to determine the unique contribution of each variable to the overall group separation.

Stepwise Selection (Exploratory Approach): As an alternative to a pre-planned step-down analysis, a stepwise procedure can be used to automatically select the dependent variables that best contribute to group separation. The procedure

enters variables one at a time, selecting the one with the maximum contribution, and then re-evaluates all variables already in the model to see if any have lost their significance and should be removed. This continues until no more variables can be entered or removed.

Example

Imagine a research study testing the effectiveness of three different teaching methods (Group A, Group B, Group C) on student performance. The researchers measure two dependent variables: Final Exam Score and Project Grade.

- Independent Variable: Teaching Method (categorical with three levels).
- Dependent Variables: Final Exam Score and Project Grade (continuous).

Instead of running two separate ANOVAs (one for each dependent variable), a MANOVA is conducted.

- Hypothesis:
  - H0: The mean vector of (Exam Score, Project Grade) is the same for all three teaching methods.
  - H1: The mean vector is different for at least one of the teaching methods.
- Results: The MANOVA shows a significant Wilks' Lambda, with an associated low p-value (e.g., $p < 0.05$).
- Interpretation: The significant MANOVA result indicates that there is a significant difference between the teaching methods when considering both the Final Exam Score and Project Grade together. The significant result is likely driven by the combined effect of the two variables, which may be correlated (e.g., students who score high on the exam also tend to have a high project grade). A follow-up analysis (like a step-down analysis or individual ANOVAs with a Bonferroni correction) would then be performed to pinpoint which specific dependent variable(s) contributed to the overall effect. For instance, it might be found that the "Project Grade" variable, when considered alongside the "Exam Score," is the primary driver of the group differences.

3 Contrasts

In MANOVA, each mean is a **vector**:

$$\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2}, \ldots, \mu_{ip})'.$$

A multivariate contrast is defined as

$$\mathbf{L} = \sum_{i=1}^{k} c_i \boldsymbol{\mu}_i, \quad \sum c_i = 0.$$

- **Example:** Suppose three groups are studied on **two outcomes** (blood pressure, cholesterol). A meaningful contrast is "treatment average vs control":

$$\mathbf{L} = \tfrac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) - \boldsymbol{\mu}_3.$$

- **Hypothesis test:**

$$H_0 : \mathbf{L} = \mathbf{0}.$$

The procedure:

1. Compute contrast-adjusted group means.
2. Form the **hypothesis SSCP matrix for the contrast**:

$$\mathbf{H}_C = \mathbf{L}'(\mathrm{Cov}(\hat{\mathbf{L}}))^{-1}\mathbf{L}.$$

3. Compare against the error SSCP $\mathbf{E}$.
4. Apply Wilks, Pillai, LH, or Roy statistics.

Interpretation: This approach focuses power on **specific scientific hypotheses** rather than the general equality of all mean vectors.

4 Tests on Individual Variables After MANOVA

After rejecting H0 in MANOVA, we often ask: which dependent variables differ?

Three strategies:

1. Univariate follow-up ANOVAs:
    - Test each dependent variable separately.
    - Problem: increases familywise error. Adjust with Bonferroni or Holm corrections.
    - Example: If MANOVA shows diet groups differ, test weight loss, cholesterol, and blood pressure individually.
2. Stepdown (Roy-Bargmann) procedure:
    - Order dependent variables (e.g., by theoretical importance).
    - Test first variable by ANOVA.
    - Next, test second variable adjusting for the first using regression residuals. Continue stepwise.
    - This controls redundancy due to correlated outcomes.
3. Canonical discriminant analysis:
    - Finds linear combinations of dependent variables that best separate groups.
    - Provides a more interpretable picture: “The main difference between treatments is a combination of high weight loss and reduced cholesterol.”

5 Two-Way Classification

With p-dimensional outcomes:

$$\mathbf{y}_{ijk} = \boldsymbol{\mu} + \boldsymbol{\alpha}_i + \boldsymbol{\beta}_j + (\boldsymbol{\alpha}\boldsymbol{\beta})_{ij} + \boldsymbol{\varepsilon}_{ijk}.$$

- Construct hypothesis matrices for each effect: $\mathbf{H}_A, \mathbf{H}_B, \mathbf{H}_{AB}$.
- Each compared to the common error matrix $\mathbf{E}$.
- Apply Wilks/Pillai/etc. for each effect.

Example: Suppose factor A = gender (male/female), factor B = treatment (drug/placebo). Outcomes = anxiety, depression scores. MANOVA tests:

- Gender effect on outcomes.
- Treatment effect.
- Gender × treatment interaction.

6 Other Models

6.1 Higher-Order Fixed Effects

With three or more factors, the model simply expands. Each main effect and interaction gets its own hypothesis matrix. Complexity lies in interpretation: higher-order interactions describe whether interaction effects depend on additional factors.

6.2 Mixed Models

When some factors are random (e.g., subjects, classrooms), expected mean squares must be derived. In MANOVA, this requires partitioning covariance correctly. Tests depend on whether a factor is fixed or random. Example: subjects nested in treatments → subject variation treated as random.

7 Checking Assumptions

1. Multivariate normality:
    - Inspect Mahalanobis distances:

      $D^2 = (y - \bar{y})^T S^{-1} (y - \bar{y})$.

      where S is the sample covariance matrix.

      Plot against chi-square quantiles.
2. Equality of covariance matrices:
    - Box's M test compares covariance matrices across groups. Large M suggests violation.
3. Independence:
    - Requires proper randomization.

Robustness: Pillai's trace is least sensitive to unequal covariances and small deviations from normality.

8 Profile Analysis

Profile analysis is used when the same set of measurements is taken across groups (e.g., repeated psychological tests). It asks three questions:

1. Parallelism: Do groups change in the same pattern across measures? (Test of interaction).

2. Level: Are average profiles elevated/depressed relative to one another? (Main effect of group).
3. Flatness: Are profiles flat (no within-group differences)? (Main effect of measure).

It is essentially a MANOVA framed to compare shapes of mean profiles.

9 Tests on a Subvector

9.1 Test for Additional Information

Partition outcomes into primary variables and new variables. Test whether new variables add information about group differences given the primary set.

9.2 Stepwise Selection of Variables

Sequentially add dependent variables to the model. At each step, test if the new variable contributes significantly to discrimination. Useful in exploratory research.

10 MANCOVA (Multivariate Analysis of Covariance)

Purpose: MANOVA ignores covariates (e.g., baseline scores). MANCOVA adjusts group comparisons for these covariates, increasing precision.

Model

$$\mathbf{y}_{ij} = \boldsymbol{\mu}_i + \mathbf{B}(\mathbf{x}_{ij} - \bar{\mathbf{x}}) + \boldsymbol{\varepsilon}_{ij}.$$

- $\mathbf{y}_{ij}$: p-dimensional outcome.
- $\mathbf{x}_{ij}$: covariates.
- $\mathbf{B}$: regression coefficient matrix.

Hypothesis

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_k \quad \text{after adjusting for covariates.}$$

Assumptions

1. Covariates linearly related to dependent variables.
2. Regression slopes homogeneous across groups.

Example

Suppose a study compares three teaching methods on math and reading scores. Covariate: pretest score. MANCOVA tests whether posttest scores differ after controlling for pretest.

Final Summary

- MANOVA extends ANOVA to multiple correlated outcomes.
- Contrasts allow specific hypotheses.
- Post-hoc univariate tests and canonical analysis clarify which variables matter.
- Extensions: two-way MANOVA, mixed models, profile analysis, repeated measures, growth curves, subvector tests.
- MANCOVA incorporates covariates to refine comparisons.