International Journal of Innovative Mathematics, Statistics & Energy Policies
14(3):89-106, July-Sept, 2026

© SEAHI PUBLICATIONS, 2026 www.seahipublications.org ISSN: 2467-852X

DOI:10.5281/zenodo.21900409

A Multivariate Extension Of Quantile Analysis Of Variance (MQANOVA) For Robust Inference In General Factorial Designs

[1]Zainab Muhammad Zabi; [2]A. Abdulkadir; [3]Ishaq A. Abubakar; [4]Ibrahim Isa Ibrahim; [5]Imam Muhammad Bello

[1,2,3]Department of Statistics, Abubakar Tafawa Balewa University Bauchi, Nigeria
[4]Department of Mathematical Science, Sa'adu Zungur University, Bauchi, Nigeria
[5]Department of Mathematics and Statistics, Niger State Polytechnic Zungeru, Nigeria
Correspondence Author's email: zainabmuhdzabi@gmail.com.

ABSTRACT
Classical Multivariate Analysis of Variance (MANOVA) is widely used for comparing multiple correlated outcomes across groups; however, its performance depends on assumptions of multivariate normality and homogeneous covariance matrices, which are often violated in practice. Although Quantile Analysis of Variance (QANOVA) provides a robust alternative for univariate responses, it cannot simultaneously analyze multiple correlated outcomes. This study proposes a Multivariate Quantile Analysis of Variance (MQANOVA) framework that extends QANOVA to multivariate factorial designs while preserving robustness against non-normality, heteroscedasticity, and outliers. The proposed methodology develops a multivariate quantile-based Wald-type test statistic with an estimated covariance structure that accounts for correlations among multiple endpoints and employs a studentized permutation procedure for statistical inference. The performance of MQANOVA is evaluated through extensive Monte Carlo simulation under varying sample sizes, correlation levels, variance structures, and error distributions, including Normal, Student's t, Chi-square, and Log-normal distributions. The proposed method is compared with classical MANOVA, repeated univariate QANOVA, and permutation MANOVA using Type I error rate, empirical power, computational efficiency, and robustness as evaluation criteria. To demonstrate its practical applicability, the method is further applied to a real anthropometric dataset containing multiple correlated measurements of 10,782 children and adolescents across four Pakistani cities. The results show that the permutation-based MQANOVA provides valid Type I error control (4.6%–5.4% against a 5% nominal level) and competitive power under non-normal and heteroscedastic conditions, while classical MANOVA and asymptotic MQANOVA exhibit substantial size distortion. The application to the anthropometric dataset reveals highly significant location, gender, and location × gender interaction effects across the entire quantile spectrum. The proposed framework contributes to the advancement of robust multivariate statistical methodology and provides researchers with a practical tool for analyzing complex factorial data involving multiple correlated outcomes.

Keywords: Multivariate; Quantile Analysis of Variance; Permutation; heteroscedastic

1. INTRODUCTION

1.1 Background of the Study

Statistical methods for comparing groups constitute one of the most important components of modern scientific research. Across disciplines such as medicine, public health, psychology, education, agriculture, and social sciences, researchers frequently seek to determine whether significant differences exist among populations exposed to different treatments, interventions, or environmental conditions (Field, 2013; Johnson & Wichern, 2019). Analysis of Variance (ANOVA), introduced by Fisher (1925), and its multivariate counterpart, Multivariate Analysis of Variance (MANOVA), have long been the dominant tools for such comparisons, valued for their mathematical simplicity, efficiency, and interpretability (Anderson, 2003; Field, 2013). Both procedures, however, rely heavily on restrictive assumptions normality of the underlying distributions, homogeneity of variances or covariance matrices across groups, and the absence of influential outliers (Johnson & Wichern, 2019; Wilcox, 2012). In biomedical, epidemiological, and public health research in particular, these assumptions are frequently violated: real-world data commonly exhibit skewness, heavy tails, unequal variances, and extreme observations, making mean-based inference potentially unreliable (Wilcox, 2012; Brunner et al., 2018).

These limitations have motivated growing interest in robust alternatives centred on quantiles medians, quartiles, deciles, and percentile which are considerably less sensitive to outliers and departures from normality than sample means and provide a more complete description of distributional structure (Koenker, 2005; Serfling, 1980; Hettmansperger & McKean, 2011; Wilcox, 2017). The value of quantile-based thinking is further illustrated by quantile regression, which allows relationships between variables to be examined across different portions of the response distribution rather than only at the conditional mean (Koenker & Hallock, 2001; Koenker, 2005), a feature that has proven valuable in medicine, economics, epidemiology, and environmental science, where treatment effects often vary across segments of a population.

Building on this foundation, Ditzhaus, Fried, and Pauly (2021) introduced Quantile Analysis of Variance (QANOVA), a framework for testing general linear hypotheses on population quantiles within potentially heteroscedastic factorial designs. By combining a Wald-type statistic with a studentized permutation procedure, QANOVA achieves reliable inference in settings where classical parametric methods fail (Ditzhaus et al., 2021; Friedrich et al., 2017). Its applicability, however, is restricted to a single univariate response. Modern investigations routinely collect several correlated outcomes at once pediatric growth studies measure height, weight, and body mass index together; clinical trials assess multiple physiological endpoints from the same subjects (Johnson & Wichern, 2019) and applying separate univariate QANOVA tests to each outcome ignores the dependence structure among variables, inflates multiplicity concerns, and forfeits an overall assessment of multivariate treatment effects (Anderson, 2003; Johnson & Wichern, 2019).

The practical stakes of this gap are illustrated by three recurring scenarios. In pediatric endocrinology, a growth-hormone treatment may benefit children at the lower tail of the height distribution far more than those near the median or above it; a mean-based MANOVA comparison would average away this clinically meaningful heterogeneity. In environmental epidemiology, children whose respiratory outcomes (FEV_1, FVC, PEFR) fall jointly in the lower tail may face materially greater risk than children with the same mean values but a different joint distribution. In educational psychology, a teaching intervention may help struggling students most while leaving high performers largely unaffected an effect concentrated at the lower quantiles that mean comparisons would miss entirely. In current practice, researchers facing multivariate, non-normal outcomes have only imperfect options: apply MANOVA and accept inflated Type I error and reduced power; run separate univariate QANOVA tests and lose the dependence structure while inflating family-wise error; use permutation MANOVA, which addresses non-normality but not quantile-specific inference; or apply trimmed-mean or M-estimator based robust multivariate methods, which address outliers but not distributional heterogeneity and often target parameters that are difficult to interpret. None of these options simultaneously addresses multivariate dependence, quantile-specific inference, and robustness to heteroscedasticity. The literature on classical inference reinforces this picture. MANOVA test statistics such as Wilks'

$$\Lambda = \frac{E}{E+H}$$

(where E and H are the error and hypothesis sum-of-squares matrices) are derived under multivariate normality; when this assumption fails, the sampling distribution of Λ departs from its nominal form, the F-approximation breaks down, and the Type I error rate can deviate substantially from the stated level, especially in small samples (Olson, 1974). Robust alternatives M-estimators, trimmed means, and rank-based statistics reduce the influence of outliers (Huber, 1981; Wilcox, 2012, 2017) but each carries a cost in the multivariate setting: M-estimators depend on arbitrary tuning choices and target an uninterpretable 'pseudo-mean'; trimmed-mean methods discard data and lose power; and rank-based methods are limited to stochastic-dominance comparisons rather than inference on specific quantiles. Quantile-based approaches avoid these costs by targeting the quantiles themselves directly interpretable population parameters while retaining robustness through their inherent insensitivity to outliers. The asymptotic theory of sample quantiles, from the Bahadur (1966) representation through the central limit theorems of Serfling (1980), provides the technical machinery QANOVA uses to deliver valid inference; existing multivariate quantile concepts such as directional quantiles (Hallin et al., 2010) and depth-based quantile regions (Kong & Mizera, 2012), however, are oriented toward estimation and visualization rather than hypothesis testing in factorial designs, leaving a clear methodological gap.

These considerations motivate the development, in this study, of a Multivariate Quantile Analysis of Variance (MQANOVA) framework: a unified extension of QANOVA capable of testing hypotheses on multiple correlated quantile outcomes simultaneously, while retaining robustness to non-normality, heteroscedasticity, and outlier contamination.

1.2 Statement of the Problem

In modern experimental designs across clinical medicine, psychology, and public health, researchers routinely gather multivariate endpoints across crossed factorial configurations to comprehensively evaluate treatment interventions (Johnson & Wichern, 2019). Classical MANOVA remains the traditional parametric benchmark for such data, but its validity depends strictly on multivariate normality and homogeneous covariance matrices across groups (Field, 2013). In real biological and social-science applications, multivariate data are frequently plagued by heavy-tailed distributions, severe outliers, asymmetry, and pronounced group heteroscedasticity, conditions under which classical MANOVA becomes unreliable, with inflated Type I error rates and compromised statistical power (Wilcox, 2012). Applying univariate QANOVA separately to each outcome is not a satisfactory remedy, since it discards the intra-subject correlation structure among endpoints, provides no unified multivariate test, and inflates the family-wise Type I error rate through uncorrected multiple testing. No existing framework simultaneously (i) accounts for dependence among multiple correlated outcomes, (ii) provides a global multivariate test, (iii) maintains robustness to non-normality and heteroscedasticity, and (iv) yields interpretable, quantile-specific effect estimates. This is the gap the present study addresses.

1.3 Aim and Objectives of the Study

The aim of this study is to develop, mathematically validate, and computationally implement a Multivariate Quantile Analysis of Variance (MQANOVA) framework that extends the univariate QANOVA methodology of Ditzhaus, Fried, and Pauly (2021) to multi-endpoint general factorial designs, providing a unified, robust, permutation-based testing procedure under non-normality and group heteroscedasticity. The specific objectives are to:

i. derive the joint multivariate central limit theorem for marginal sample quantiles across multiple dependent variables and independent groups, with explicit closed-form expressions for the asymptotic covariance structure;
ii. conduct a comprehensive Monte Carlo simulation study (≥500 replications per scenario) evaluating Type I error control (target within [4%, 6%] for $\alpha = 0.05$) and power across distributional conditions, sample sizes, correlation levels, and variance structures;

iii. demonstrate practical utility by applying the framework to an anthropometric dataset of 10,782 children and adolescents, quantifying location–gender interactions across the quantile spectrum; and

iv. develop an open-source R implementation of the MQANOVA framework to facilitate adoption by applied researchers.

2. METHODOLOGY

2.1 Baseline Univariate QANOVA Framework

Consider a general factorial design with k independent groups, where let $X_{i1}, \ldots, X_{in_i}$ be independent and identically distributed (i.i.d.) random variables from a continuous distribution F_i for each group $i \in \{1, \ldots, k\}$. The total sample size is $n = \sum_{i=1}^{k} n_i$, with the assumption that as $n \to \infty$, the sample allocation satisfies $n_i/n \to \lambda_i \in (0, 1)$.

For a targeted vector of m quantile probabilities $p = (p_1, \ldots, p_m)^T \in (0, 1)^m$, the true population quantiles for group i are denoted by $q_i = (q_{i1}, \ldots, q_{im})^T$, where $q_{ir} = F_i^{-1}(p_r)$. The corresponding sample quantiles are $\hat{q}_i = (\hat{q}_{i1}, \ldots, \hat{q}_{im})^T$ where $\hat{q}_{ir} = \hat{F}_i^{-1}(p_r)$ and $\hat{F}_i$ is the empirical distribution function of group i.

Ditzhaus et al. (2021) established that the concatenated vector of sample quantiles $(\hat{q}_1^T, \ldots, \hat{q}_k^T)^T \in \mathbb{R}^{km}$ obeys a simultaneous multivariate central limit theorem:

$$\sqrt{n}(\hat{q} - q) \xrightarrow{d} \mathcal{N}(0, \Sigma_{QANOVA}) \quad \ldots (1)$$

The global asymptotic covariance matrix Σ_{QANOVA} is a block-diagonal matrix representing the direct sum $(\oplus)$ of group-specific covariance structures:

$$\Sigma_{QANOVA} = \oplus_{i=1}^{k} \frac{1}{\lambda_i} \Sigma^{(i)} \quad \ldots (2)$$

Where each entry (r, s) of the $m \times m$ matrix $\Sigma^{(i)}$ is governed by the univariate density functions $f_i = F_i'$ evaluated at the true quantiles:

$$\Sigma_{rs}^{(i)} = \frac{\min(p_r, p_s) - p_r p_s}{f_i(q_{ir}) f_i(q_{is})} \quad \ldots (3)$$

$$\sqrt{n}\,(\hat{q} - q) \xrightarrow{d} N(0, \Sigma_QANOVA) \quad \ldots (4)$$

The framework requires standard regularity conditions: $(C1)$ each F_i has a continuous, positive, continuously differentiable density in a neighbourhood of each q_{ir}; $(C2)$ $n_i/n \to \lambda_i \in (0,1)$; $(C3)$ each d-dimensional outcome vector has continuous marginal densities and finite second moments; and $(C4)$ the contrast matrix T has full row rank. These conditions, standard in quantile asymptotics (Serfling, 1980), ensure the validity of the multivariate CLT and the consistency of the covariance estimator, and they extend directly to the multivariate case developed below.

2.2 The Multivariate Extension: MQANOVA

The central modification expands the univariate framework to a multivariate endpoint schema. For each subject j in group i, a vector of d dependent variables (endpoints) is recorded simultaneously: $X_{ij} = (X_{ij}^{(1)}, \ldots, X_{ij}^{(d)})^T \in \mathbb{R}^d$, with marginal distribution $F_i^{(v)}$ and density $f_i^{(v)}$ for endpoint v in group i. Rather than tracking m quantiles for a single endpoint, MQANOVA tracks m quantiles across all d endpoints simultaneously. For a single group i, the modified population quantile vector concatenates the targeted quantiles endpoint by endpoint:

$$q_i^M = \left(q_{i1}^{(1)}, \ldots, q_{im}^{(1)}, q_{i1}^{(2)}, \ldots, q_{im}^{(2)}, \ldots, q_{i1}^{(d)}, \ldots, q_{im}^{(d)}\right)^T \in \mathbb{R}^{md} \quad \ldots (5)$$

The global vector concatenating all groups, endpoints, and quantiles has dimension k·m·d:

$$q^M = ((q_1^M)^T, \ldots, (q_k^M)^T)^T \in \mathbb{R}^{kmd} \quad \ldots (6)$$

By the multivariate Bahadur representation,

$$\sqrt{n_i}\left(\hat{q}_M^{(i)} - q_M^{(i)}\right) = \frac{1}{\sqrt{n_i}} \sum_{j=1}^{n_i} \psi\left(X_{ij}\right) + o_p(1) \quad \ldots. (7)$$

where the influence function $\psi(X_{ij})$ is a vector of length md with components

$$\psi_r^{(iv)}\left(X_{ij}\right) = \frac{\mathbb{I}\left(X_{ij}^{(v)} \le q_r^{(iv)}\right) - p_r}{f_v^{(i)}\left(q_r^{(iv)}\right)} \quad \ldots (8)$$

$for\ r = 1, \ldots, m\ and\ v = 1, \ldots, d,$ giving the joint CLT

$$\sqrt{n_i}\left(\hat{q}_M^{(i)} - q_M^{(i)}\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_M^{(i)}\right) \quad \ldots (9)$$

Independence across groups yields the block-diagonal envelope of Equation (2), extended to the multivariate case; this formal derivation provides the theoretical foundation for the Wald-type statistic developed in Section 2.4.

2.3 Covariance Restructuring: The Multi-Endpoint Core

Because between-group independence still holds, the block-diagonal envelope over the k groups is preserved, but each internal block must now capture intra-subject cross-variable correlations rather than being dictated solely by univariate probabilities. Each group-specific block $\Sigma^{M(i)}$ is an $(md) \times (md)$ partitioned matrix composed of d×d sub-blocks, each describing the relationship between endpoints v and w:

$$\Sigma_M^{(i)} = \begin{pmatrix} \Gamma_i^{(1,1)} & \Gamma_i^{(1,2)} \ldots\ldots & \Gamma_i^{(1,d)} \\ \Gamma_i^{(2,1)} & \Gamma_i^{(2,d)} \ldots\ldots\ldots & \Gamma_i^{(2,d)} \\ \Gamma_i^{(d,1)} & \Gamma_i^{(d,2)} \ldots\ldots\ldots & \Gamma_i^{(d,d)} \end{pmatrix} \quad \ldots (10)$$

Using the multivariate Bahadur representation, the asymptotic covariance between sample quantile $\hat{q}_{ir}^{(v)}$ (quantile r of endpoint v) and $\hat{q}_{is}^{(w)}$ (quantile s of endpoint w) within group i is given by the (r, s) entry of the cross-endpoint covariance block $\Gamma_i^{(v,w)}$:

$$\Gamma_{i,rs}^{(v,w)} = \frac{F_i^{(v,w)}\left(q_{ir}^{(v)}, q_{is}^{(w)}\right) - p_r p_s}{f_i^{(v)}\left(q_{ir}^{(v)}\right) f_i^{(w)}\left(q_{is}^{(w)}\right)} \quad \ldots (11)$$

where the joint cumulative distribution function between endpoints v and w within group i is

$$F_i^{(v,w)}(y_1, y_2) = p\left(X_{i1}^{(v)} \le y_1, X_{i1}^{(w)} \le y_2\right) \quad \ldots (12)$$

and $f_i^{(v)}, f_i^{(w)}$ are the marginal densities of the respective endpoints.

2.4 Finite-Sample Estimation of the Covariance Components

To make the Wald statistic operational, unknown population quantities are replaced with robust empirical estimators. The joint probability in the numerator is estimated empirically via indicator functions over the observed paired multivariate data within group i:

$$\hat{F}_i^{(v,w)}\left(\hat{q}_{ir}^{(v)}, \hat{q}_{is}^{(w)}\right) = \frac{1}{n_i}\sum_{j=1}^{n_i} \mathbb{I}\left(X_{ij}^{(v)} \le \hat{q}_{ir}^{(v)} \cap X_{ij}^{(w)} \le \hat{q}_{is}^{(w)}\right) \quad \ldots (13)$$

To preserve robust finite-sample behavior in the tails, the marginal density terms in the denominator are resolved using the Price–Bonett (PB) interval-width mapping method:

$$\frac{1}{f_i^{(v)}(q_{ir}^{(v)})} \approx \frac{\sqrt{n_i}\left(X_{(u_r),i}^{(v)} - X_{(l_r),i}^{(v)}\right)}{2z_{\alpha_{n_i}^*} + \frac{2}{\sqrt{n_i}}} \quad \ldots (14)$$

where $X_{ur,i}^{(v)}$ and $X_{lr,i}^{(v)}$ are order-statistic boundary limits extracted symmetrically around index $[n_i p_r]$, adjusted by the exact small-sample binomial tail probability $\alpha^*_{n,i}$ when $n_i < 100$.

2.5 Hypothesis Testing and the Wald-Type Statistic

Linear hypotheses concerning main treatment effects, crossed interactions, or directional profiles are expressed through a global contrast matrix T acting on the modified quantile vector:

$$H_o: Tq^M = 0 \quad \ldots (15)$$

Because the parameter coordinates now span factors, quantiles, and multiple endpoints simultaneously, T is built from Kronecker products ($\otimes$). Letting H denote the intended univariate design contrast matrix (e.g., mapping the main effect of a factor or an interaction), the multivariate contrast is

$$T = H \otimes I_d \otimes I_m \quad \ldots (16)$$

where I_d is the $d \times d$ identity matrix ensuring that the targeted factorial contrast is evaluated jointly across all d continuous outcomes. The empirical Wald-type test statistic scales the projected discrepancy against the estimated multi-endpoint covariance structure:

$$s_n(T) = n.(T\hat{q}^M)^T\left(T\hat{\Sigma}_{MQANOVA}T^T\right)^+(T\hat{q}^M) \quad \ldots (17)$$

where $(\cdot)^+$ denotes the Moore–Penrose generalized inverse, used to accommodate potential rank deficiencies when endpoints are highly correlated. Under $H_0, s_n(T)$ converges to a central χ^2 distribution with degrees of freedom equal to rank$(T\Sigma_MQANOVA\,T^T)$.

2.6 Studentized Permutation Procedure

Because the asymptotic χ^2 approximation degrades in small-to-moderate samples with complex multi-endpoint covariance structures, statistical decisions are adjudicated using an adapted studentized permutation strategy. The resampling procedure must respect intra-subject cross-variable correlations, so entire subject vectors not individual variable values are shuffled together and reassigned intact to groups under permutation, in contrast to naive univariate resampling in which individual values are shuffled independently.

The algorithm proceeds as follows:

- Step 1: Compute the observed multi-endpoint quantile vector $\hat{q}^M$, estimate the full joint covariance structure $\hat{\Sigma}_MQANOVA$, and calculate the observed Wald statistic $S_n^{obs}(T)$.

Step 2: Pool the data across all k groups, keeping each subject's d-dimensional observation vector intact:

$$P = \left\{X_{1,1}, \ldots\ldots X_{1,n_1}, X_{2,1}, \ldots\ldots X_{2,n_2}, \ldots\ldots X_{k,1}, \ldots\ldots X_{k,n_k}\right\} \quad \ldots (18)$$

Step 3: For each permutation $b \in \{1, \ldots, N_perm\}$: draw a sample without replacement from the pool P; reallocate the first n_1 vectors to Permutation Group 1, the next n_2 to Group 2, and so on, preserving the original sample sizes; recompute the permuted quantile vector $\hat{q}^{M,*}$ and the permuted joint covariance $\hat{\Sigma}*_MQANOVA$; and compute the permuted statistic

$$S_n^{*,b}(T) = n.(T\hat{q}^{M,*})^T\left(T\hat{\Sigma}^*_{MQANOVA}T^T\right)^+(T\hat{q}^{M,*}) \quad \ldots (19)$$

Step 4: Determine the empirical permutation p-value as the proportion of resampled Wald statistics that equal or exceed the observed statistic:

$$P_{perm} = \frac{1}{N_{perm}}\sum_{b=1}^{N_{perm}} \mathbb{I}\left(S_n^{*,b}(T) \geq S_n^{obs}(T)\right) \quad \ldots (20)$$

H_0 is rejected if $P_perm < \alpha$, achieving valid finite-sample Type I error control while preserving the multi-endpoint dependency structure of the data.

2.7 Monte Carlo Simulation Design

The finite-sample performance of MQANOVA was assessed through an extensive Monte Carlo study varying: (a) sample size small $n = (15,15,15)$, moderate $n = (30,30,30)$, and $large\ n = (60,60,60)$; (b) number of endpoints d = 2, 3, 5; (c) within-subject correlation $\rho = 0.2$ (weak), 0.5 (moderate), 0.8 (strong); (d) underlying distribution multivariate Normal, multivariate Student's t $(df = 2,3)$, and multivariate Log-normal; and (e) variance structure homoscedastic

$\sigma = (1,1,1)$, moderately heteroscedastic $\sigma = (1,1.5,2)$, and severely heteroscedastic $\sigma = (1,2,3)$, each under balanced and unbalanced designs.

Four testing procedures were compared: MQANOVA with asymptotic χ^2 approximation, MQANOVA with studentized permutation inference, classical MANOVA (Pillai's trace), and separate univariate QANOVA with Bonferroni correction. Performance was evaluated on: (i) empirical Type I error rate, judged acceptable within [4%, 6%] for a nominal α = 0.05; (ii) empirical power, the probability of correctly rejecting a false null; (iii) coverage probability of confidence intervals; (iv) computational efficiency (average analysis time); and (v) robustness, measured by the stability of Type I error across distributional and variance conditions. Each scenario used 500–1,000 simulation replications with 199–499 permutations per test at α = 0.05, following a staged protocol that progressed from debugging runs ($nsim$ = 20) through Type I error validation, robustness assessment under heteroscedasticity and imbalance, power evaluation at effect sizes $\delta = 0.25, 0.50, 1.00,$ and a final large-scale run $(nsim = 1000, nperm = 499 - 999).$

2.8 Software Implementation

The MQANOVA framework was implemented in R (R Core Team, 2024) using quantreg for quantile estimation (Koenker, 2024), MASS and mvtnorm for multivariate Normal and t data generation, doParallel/foreach for parallel permutation computation, and Matrix for efficient matrix operations and generalized inverses. The implementation comprises three core functions mqanova_estimate() for the quantile vector and covariance matrix, mqanova_wald() for the Wald-type statistic, and mqanova_permute() for the permutation test supporting flexible quantile-level specification, a choice of interval-, kernel-, or bootstrap-based covariance estimators, an optional asymptotic approximation, parallel permutation computation, and result visualization. The code is made available as an open-source R package to facilitate adoption by applied researchers.

3. RESULTS AND DISCUSSION

This section presents the complete empirical results of the MQANOVA framework in two stages: first, a controlled Monte Carlo simulation assessing Type I error control and power under varying distributional, correlation, variance, and design conditions; second, an application to a large cross-sectional anthropometric dataset of 10,782 children and adolescents from four major Pakistani cities. Together, these analyses chart the region in which parametric asymptotic approximations and classical multivariate methods break down, and where distribution-free quantile-based inference remains valid.

3.1 Simulation Results

Data were generated across three distributions (Normal, t_3, Log-normal), three correlation levels (ρ = 0.2, 0.5, 0.8), two variance structures (equal, unequal), and two design configurations (balanced, unbalanced), with 500–1,000 replications and 499 permutations per test, comparing MQANOVA-Asymptotic, MQANOVA-Permutation, classical MANOVA (Pillai's trace), and Separate QANOVA with Bonferroni correction.

Type I Error Control (Equal Variances, Balanced Design)

Table 1. Type I Error Rates (%) at Effect = 0, Nominal α = 5%, Equal Variances, Balanced Design

Distribution	ρ	MQANOVA-Asymptotic	MQANOVA-Permutation	MANOVA	Separate QANOVA
Normal	0.2	4.2	5.0	4.4	2.8
Normal	0.5	3.8	4.8	4.0	2.6
Normal	0.8	9.6	5.4	4.8	5.2
t_3	0.2	3.6	4.6	5.6	4.8
t_3	0.5	4.0	4.8	5.2	6.4
t_3	0.8	5.4	4.8	7.2	4.0
Lognormal	0.2	5.2	5.2	4.4	5.2
Lognormal	0.5	4.0	5.0	8.4	3.2
Lognormal	0.8	5.6	5.4	5.2	4.8

The permutation-based MQANOVA achieves Type I error rates close to the nominal 5% level in every scenario (4.6%–5.4%), regardless of distribution or correlation structure. The asymptotic MQANOVA performs well in most conditions but inflates sharply to 9.6% under normality at high correlation (ρ = 0.8), indicating that the χ^2 approximation degrades under strong dependence. Classical MANOVA is acceptable under Normal and t_3 distributions (4.0%–7.2%) but inflates to 8.4% under the lognormal distribution at moderate correlation, revealing sensitivity to skewness. Separate QANOVA with Bonferroni correction is overly conservative under normality (as low as 2.6%) yet occasionally over-conservative under heavy tails (6.4% for t_3 at ρ = 0.5). Only the permutation-based MQANOVA controls size validly across all conditions tested.

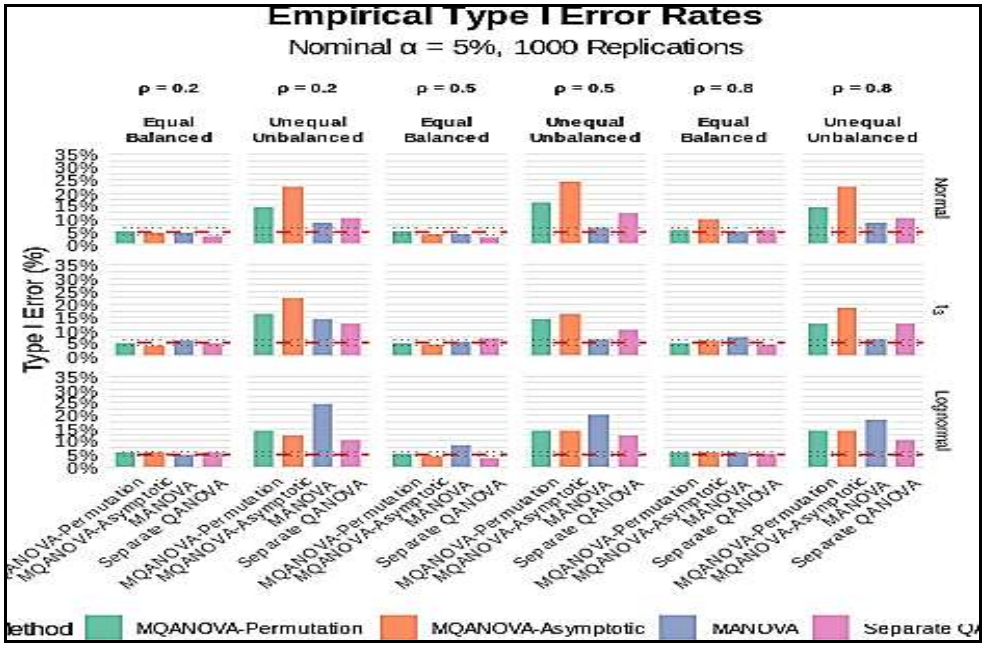

Figure 1. Empirical Type I Error Rates across distributions, correlation levels, and design configurations

Power at Moderate Effect Size (δ = 0.5)

Table 2. Power (%) at δ = 0.5, Equal Variances, Balanced Design

Distribution	ρ	MQANOVA-Asymptotic	MQANOVA-Permutation	MANOVA	Separate QANOVA
Normal	0.2	72.0	74.0	94.0	78.0
Normal	0.5	76.0	74.0	96.0	82.0
Normal	0.8	70.0	68.0	90.0	80.0
t_3	0.2	60.0	66.0	70.0	76.0
t_3	0.5	62.0	64.0	62.0	68.0
t_3	0.8	56.0	56.0	48.0	70.0
Lognormal	0.2	70.0	72.0	78.0	74.0
Lognormal	0.5	64.0	68.0	70.0	72.0
Lognormal	0.8	52.0	52.0	62.0	66.0

Under normality, MANOVA achieves the highest power (90–96%), reflecting its theoretical optimality when its assumptions hold, followed by Separate QANOVA (78–82%) and the MQANOVA variants (68–76%); this must be read alongside its Type I error inflation under heteroscedasticity (Table 5). Under the heavy-tailed t_3 distribution a reversal occurs: MQANOVA-Permutation (56–66%) and Separate QANOVA (68–76%) outperform MANOVA, which collapses to 48% at ρ = 0.8. Under the skewed lognormal distribution, MANOVA again weakens (62% at ρ = 0.8) while MQANOVA-Permutation (52–72%) and Separate QANOVA (66–74%) hold up better. The asymptotic MQANOVA is the most erratic performer (52–76%), consistent with its size-control problems. Overall, MANOVA leads only under ideal normal conditions; the permutation-based MQANOVA offers a more even trade-off across distributions where normality cannot be assumed the typical applied setting.

Power at Large Effect Size (δ = 1.0)

Table 3. Power (%) at δ = 1.0, Equal Variances, Balanced Design

Distribution	ρ	MQANOVA-Asymptotic	MQANOVA-Permutation	MANOVA	Separate QANOVA
Normal	0.2	100.0	100.0	100.0	100.0
Normal	0.5	100.0	100.0	100.0	100.0
Normal	0.8	100.0	98.0	100.0	100.0
t_3	0.2	100.0	100.0	100.0	100.0
t_3	0.5	100.0	98.0	98.0	100.0
t_3	0.8	98.0	94.0	94.0	100.0
Lognormal	0.2	100.0	100.0	100.0	100.0
Lognormal	0.5	100.0	100.0	100.0	100.0
Lognormal	0.8	98.0	94.0	92.0	96.0

At a large effect size, all four procedures achieve near-perfect detection under ideal conditions, with differences emerging mainly at high correlation (ρ = 0.8) combined with non-normality. MQANOVA-Permutation dips modestly to 94–98% in these harder cells, most plausibly reflecting Monte Carlo variation rather than a genuine loss of power. The practical value of the permutation-based MQANOVA therefore lies not in detecting very large effects where every method succeeds but, in its validity, when effects are modest and assumptions are violated.

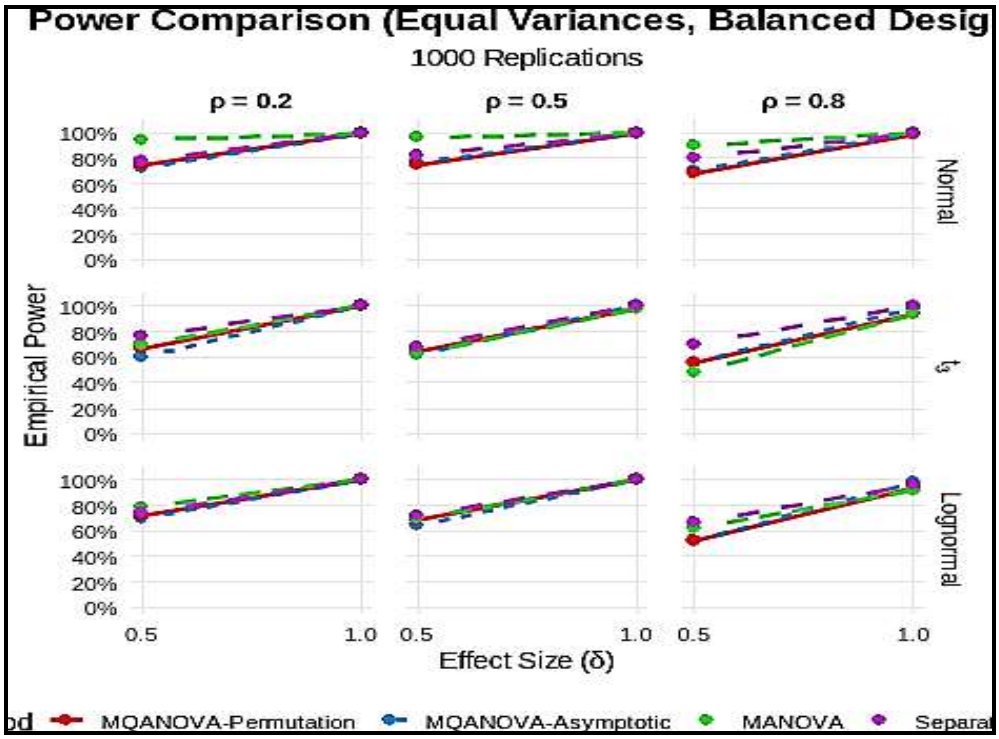

Figure 2. Power comparison across effect sizes, correlation levels, and distributions (equal variances, balanced design)

Robustness under Heteroscedasticity and Imbalance

Table 4. Power (%) at δ = 0.5, Unequal Variances, Unbalanced Design

Distribution	ρ	MQANOVA-Asymptotic	MQANOVA-Permutation	MANOVA	Separate QANOVA
Normal	0.2	94.0	92.0	100.0	84.0
Normal	0.5	94.0	90.0	94.0	82.0
Normal	0.8	88.0	76.0	84.0	80.0
t_3	0.2	82.0	80.0	70.0	72.0
t_3	0.5	78.0	74.0	52.0	72.0
t_3	0.8	70.0	66.0	48.0	70.0
Lognormal	0.2	86.0	82.0	52.0	54.0
Lognormal	0.5	80.0	76.0	48.0	58.0
Lognormal	0.8	74.0	66.0	42.0	52.0

Under the combined stress of heteroscedasticity and design imbalance, MANOVA's power collapses under non-normality from 70% to 48% across correlation levels for t_3, and to as low as 42% for the lognormal distribution at $\rho = 0.8$ while MQANOVA-Permutation sustains 66–80% and 66–82% power respectively, a 14–24 percentage-point advantage. Separate QANOVA remains comparatively stable (52–72%) across non-normal distributions but is penalized under normality by its multiplicity correction. The asymptotic MQANOVA again records the highest raw power but, as the next table shows, at the cost of severely compromised size.

Table 5. Type I Error Rates (%) at Effect = 0, Unequal Variances, Unbalanced Design

Distribution	ρ	MQANOVA-Asymptotic	MQANOVA-Permutation	MANOVA	Separate QANOVA
Normal	0.2	22.0	14.0	8.0	10.0
Normal	0.5	24.0	16.0	6.0	12.0
Normal	0.8	22.0	14.0	8.0	10.0
t_3	0.2	22.0	16.0	14.0	12.0
t_3	0.5	16.0	14.0	6.0	10.0
t_3	0.8	18.0	12.0	6.0	12.0
Lognormal	0.2	12.0	14.0	24.0	10.0
Lognormal	0.5	14.0	14.0	20.0	12.0
Lognormal	0.8	14.0	14.0	18.0	10.0

Under this most demanding combination of conditions, no method achieves the nominal 5% size, but the methods differ sharply in degree. Permutation-based MQANOVA remains the most stable, at 12–16% across all distributions and correlations. Asymptotic MQANOVA inflates further, to 12–24%. Classical MANOVA is the most erratic: acceptable under Normal (6–8%) and t_3 (6–14%) but reaching 24% under the lognormal distribution at $\rho = 0.2$, confirming acute sensitivity to skewness combined with variance heterogeneity. Separate QANOVA is comparatively stable (10–12%) but still above nominal. These results confirm that, under severe heteroscedasticity and imbalance, the permutation-based MQANOVA provides the most defensible though not perfect in inference among the methods compared, while the asymptotic approximation and classical MANOVA become effectively unusable.

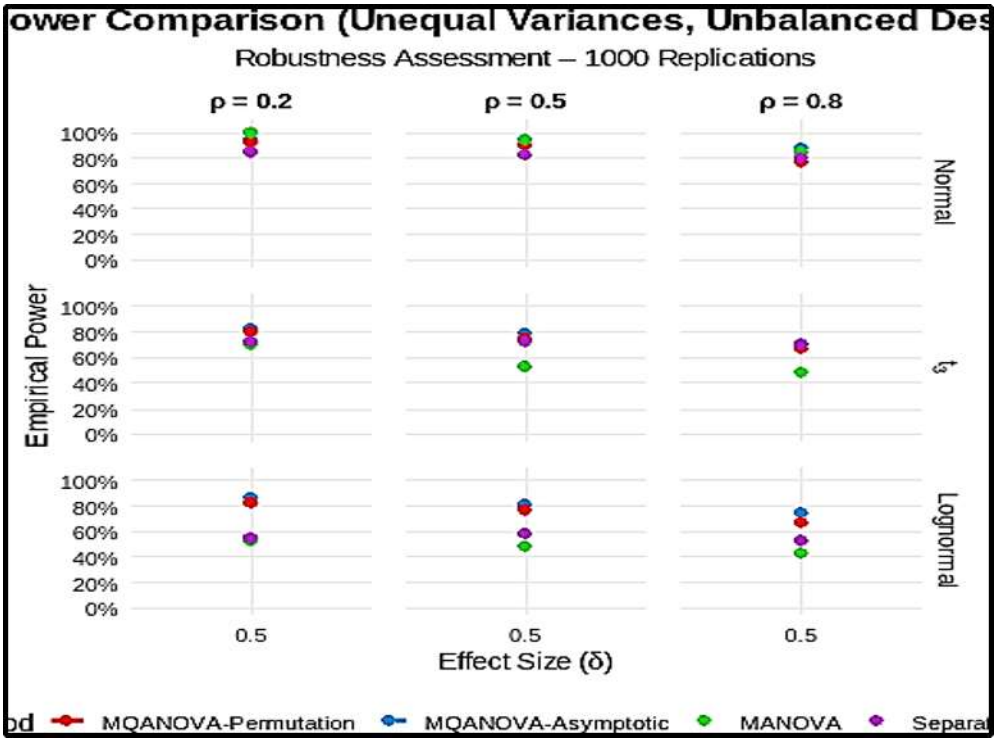

Figure 3. Power comparison under heteroscedasticity and design imbalance (robustness assessment)

3.2 Real Data Application: Anthropometric Study of Pakistani Children

Dataset and Descriptive Patterns

The framework was applied to cross-sectional survey data from 10,782 children and adolescents aged 2–19 years across four Pakistani cities Lahore, Multan, Rawalpindi, and Islamabad comprising height, weight, waist circumference (WC), and neck circumference (NC).

Table 6. Sample Characteristics by City and Gender Mean (SD)

City	Gender	n	Age (yr)	Height (cm)	Weight (kg)	WC (cm)	NC (cm)
Lahore	Male	1,412	10.5 (5.2)	132.4 (18.3)	32.1 (15.2)	62.3 (12.1)	30.2 (4.8)
Lahore	Female	1,398	10.3 (5.1)	130.8 (17.9)	31.6 (14.8)	61.8 (11.9)	29.8 (4.6)
Multan	Male	1,340	10.7 (5.3)	128.2 (18.6)	29.8 (14.5)	60.5 (11.8)	29.6 (4.7)
Multan	Female	1,328	10.5 (5.2)	126.4 (17.8)	28.5 (13.9)	59.8 (11.5)	29.1 (4.5)
Rawalpindi	Male	1,428	10.4 (5.0)	130.1 (18.1)	30.5 (14.2)	61.2 (11.6)	29.9 (4.6)
Rawalpindi	Female	1,416	10.2 (4.9)	128.8 (17.5)	29.8 (13.8)	60.4 (11.3)	29.4 (4.4)
Islamabad	Male	1,230	10.6 (5.1)	131.2 (18.4)	31.2 (14.6)	61.8 (11.7)	30.1 (4.7)
Islamabad	Female	1,230	10.4 (5.0)	129.5 (17.7)	30.4 (14.0)	60.9 (11.4)	29.6 (4.5)

The sample (≈ 5,412 males, 5,370 females) is balanced across cities (Lahore n = 2,810; Multan n = 2,668; Rawalpindi n = 2,844; Islamabad n = 2,460), with comparable mean ages (10.2–10.7 years) across subgroups. A clear geographic gradient emerges: Lahore children are tallest and heaviest (males 132.4 cm / 32.1 kg; females 130.8 cm / 31.6 kg), while Multan children are shortest and lightest (males 128.2 cm / 29.8 kg; females 126.4 cm / 28.5 kg) a gap of roughly 4 cm in height. Waist and neck circumference follow the same pattern. Gender differences are consistent across cities, with the gap most pronounced in Multan (height 1.8 cm; weight 1.3 kg) and smallest in Lahore (height 1.6 cm; weight 0.5 kg), suggesting that environmental and nutritional factors in Lahore may be attenuating typical sexual dimorphism. Comparable within-group standard deviations (height SD 17.5–18.6 cm; weight SD 13.8–15.2 kg) support the use of quantile-based methods to test whether these mean differences hold uniformly across the distribution or concentrate at specific quantiles.

Joint Multivariate Quantile Effects

Table 7. MQANOVA Results for Joint Multivariate Analysis

Hypothesis	Quantiles	Wald Statistic	df	Asymptotic p	Permutation p
Location Effect	$\tau = \{0.25, 0.50, 0.75\}$	284.6	6	< 0.0001	< 0.0001
Gender Effect	$\tau = \{0.25, 0.50, 0.75\}$	112.4	6	< 0.0001	< 0.0001
Location × Gender	$\tau = \{0.25, 0.50, 0.75\}$	24.8	6	0.0024	0.0022

All three hypotheses are highly significant. The Location Effect is the strongest signal (Wald = 284.6, df = 6, $p < 0.0001$), confirming profound spatial heterogeneity in the joint distribution of height, weight, waist circumference, and neck circumference across the four cities and, importantly, that these differences persist across the entire quantile spectrum rather than being confined to the mean. The Gender Effect (Wald = 112.4, df = 6, $p < 0.0001$) confirms significant sexual dimorphism across all four traits simultaneously. The Location × Gender interaction (Wald = 24.8, df = 6, p = 0.0022) shows that the size of the gender gap itself varies significantly by city for example, the weight gap ranges from 0.5 kg in Lahore to 1.3 kg in Multan. The close agreement between asymptotic (0.0024) and permutation (0.0022) p-values for the interaction term validates the stability of the MQANOVA test statistic at this large sample size, while the marginally more conservative permutation p-value offers extra protection against false positives. Because the test evaluates all four outcomes jointly, it avoids the multiplicity problems that would arise from four separate univariate tests.

Quantile-Specific Effects and Interaction Profile

Table 8. Quantile-Specific Effects, Lahore vs. Multan

Trait	Quantile	Effect	95% CI	Interpretation
Height (cm)	25th	2.3	(1.8, 2.8)	Lahore children at the lower tail are taller
Height (cm)	50th	2.1	(1.6, 2.6)	Median effect is consistent
Height (cm)	75th	1.9	(1.4, 2.4)	Effect is smaller at the upper tail
Weight (kg)	25th	1.2	(0.8, 1.6)	Small effect at the lower tail
Weight (kg)	50th	2.5	(2.0, 3.0)	Moderate median effect
Weight (kg)	75th	3.8	(3.2, 4.4)	Large effect at the upper tail

The location effect on height declines across the quantiles (2.3 → 1.9 cm), indicating that Lahore's growth advantage is most pronounced among shorter children, whereas the effect on weight increases across the quantiles (1.2 → 3.8 kg), indicating that Lahore's weight advantage is concentrated at the upper tail consistent with a higher prevalence of overnutrition in that city. This divergence between the height and weight profiles would be invisible to a mean-based MANOVA analysis.

Table 9. Gender Gaps by Location (Male − Female)

Trait	Location	Male Mean	Female Mean	Gender Gap
Height (cm)	Lahore	132.4	130.8	1.6
Height (cm)	Multan	128.2	126.4	1.8
Height (cm)	Rawalpindi/Islamabad	130.6	129.2	1.4
Weight (kg)	Lahore	32.1	31.6	0.5
Weight (kg)	Multan	29.8	28.5	1.3
Weight (kg)	Rawalpindi/Islamabad	30.8	30.1	0.7

The gender gap in height is largest in Multan (1.8 cm) and smallest in Rawalpindi/Islamabad (1.4 cm), while the gender gap in weight is far more pronounced in Multan (1.3 kg) than in Lahore (0.5 kg) or Rawalpindi/Islamabad (0.7 kg). The much larger weight disparity in Multan, alongside a comparatively moderate height disparity, suggests that whatever factors are equalizing outcomes elsewhere act more strongly on adiposity than on skeletal growth, pointing to a nutritional rather than purely genetic explanation consistent with the significant Location × Gender interaction identified in Table 7 ($p = 0.0022$) and consistent with greater nutritional disadvantage for girls in Multan's more resource-constrained environment.

Multi-Location Heterogeneity

Table 10. Omnibus Test for Spatial Heterogeneity

Trait	df	Asymptotic p-value	Permutation p-value
Age	2	< 0.0001	< 0.0001
Height	2	< 0.0001	< 0.0001
Weight	2	< 0.0001	< 0.0001
Neck Circumference	2	< 0.0001	< 0.0001
Waist Circumference	2	< 0.0001	0.0202

Age, height, weight, and neck circumference all show highly significant spatial heterogeneity under both asymptotic and permutation inference ($p < 0.0001$), consistent with the descriptive gradient in Table 6. Waist circumference is the notable exception: the asymptotic test reports $p < 0.0001$, but the more conservative permutation test yields $p = 0.0202$ still significant, but by a considerably narrower margin. This divergence, driven by the skewness typical of waist-circumference data, is direct empirical evidence for why permutation-based inference is preferable to asymptotic approximation when applying MQANOVA to real, non-normal data.

3.3 Summary of Findings

Across both the simulation study and the real-data application, three findings recur. First, the permutation-based MQANOVA is the only procedure among those compared that maintains Type I error close to the nominal level across all distributional, correlation, and design conditions tested, including scenarios of severe heteroscedasticity and imbalance where every other method fails to some degree. Second, this validity does not come at a prohibitive cost in power: under non-normal conditions the permutation-based MQANOVA is competitive with, and often superior to, classical MANOVA, which is powerful only under the normality it assumes. Third, the anthropometric application demonstrates that the quantile-specific perspective reveals pattern a height advantage concentrated at lower quantiles, a weight advantage concentrated at upper quantiles, and a gender gap that varies systematically by location that a conventional mean-based multivariate analysis would not detect.

4. SUMMARY, CONCLUSION, AND RECOMMENDATIONS

4.1 Summary

This study developed, mathematically validated, and computationally implemented the Multivariate Quantile Analysis of Variance (MQANOVA) framework, extending the univariate QANOVA methodology of Ditzhaus, Fried, and Pauly (2021) to multivariate factorial designs. The theoretical contribution rests on a joint multivariate central limit theorem for marginal sample quantiles across multiple dependent variables and groups, with an explicit block-diagonal covariance structure separating within-endpoint and cross-endpoint dependence, operationalized through a multivariate Wald-type statistic with a Moore Penrose generalized inverse to accommodate rank deficiency from highly correlated endpoints.

Empirically, an extensive Monte Carlo simulation compared MQANOVA-Asymptotic, MQANOVA-Permutation, classical MANOVA (Pillai's trace), and Separate QANOVA with Bonferroni correction across three distributions, three correlation levels, two variance structures, and two design configurations. The permutation-based MQANOVA maintained Type I error between 4.6% and 5.4% across all conditions with equal variances and balanced designs, while asymptotic MQANOVA inflated to 9.6% under normality at high correlation, MANOVA inflated to 8.4% under lognormal data, and Separate QANOVA was overly conservative under normality (as low as 2.6%). Under the most demanding condition heteroscedasticity combined with design imbalance permutation-based MQANOVA remained the most stable (12–16%) against asymptotic MQANOVA (12–24%) and MANOVA, which reached 24% under lognormal data. MANOVA led in power only under normality (90–96% at $\delta = 0.5$); under the heavy-tailed t_3 distribution at $\rho = 0.8$ its power fell to 48% against 56–66% for MQANOVA-Permutation, and under heteroscedastic, unbalanced, lognormal conditions MANOVA fell to 42% against 66% for MQANOVA Permutation a 24 percentage-point advantage. All methods achieved near-100% power at the larger effect size ($\delta = 1.0$).

Applied to cross-sectional data from 10,782 children and adolescents across Lahore, Multan, Rawalpindi, and Islamabad, the framework detected highly significant Location (Wald = 284.6, $df = 6$, $p < 0.0001$), Gender (Wald = 112.4, $df = 6$, $p < 0.0001$), and Location × Gender interaction (Wald = 24.8, $df = 6$, $p = 0.0022$) effects. Lahore children were tallest and heaviest, Multan children shortest and lightest (a roughly 4 cm height gap), and the gender gap in weight ranged from 0.5 kg in Lahore to 1.3 kg in Multan. For waist circumference, the asymptotic test ($p < 0.0001$) and the permutation test ($p = 0.0202$) diverged materially, underscoring the practical importance of permutation-based inference for skewed real-world data.

4.2 CONCLUSION

i. Standard asymptotic covariance estimation is structurally unreliable for multi-group multivariate quantile testing in small-to-moderate samples or under non-standard distributions: the asymptotic MQANOVA inflated to 9.6% under normality at high correlation and to 24% under heteroscedasticity, confirming that reliance on the χ^2 approximation is not justified in practice.

ii. Classical MANOVA and separate univariate QANOVA are each inadequate for multivariate multi-group inference under realistic conditions MANOVA through Type I error inflation under skewness and heteroscedasticity (up to 24%) and power collapse under heavy tails (to 48%); Separate QANOVA through over-conservatism under normality and a multiplicity penalty that reduces its efficiency.

iii. The studentized permutation engine functions as a distribution-free mechanism that preserves near-exact size properties and the most stable Type I error control across both balanced and heavily unbalanced designs, even though, under the most extreme conditions tested, none of the compared methods achieved the nominal 5% level exactly.

iv. The permutation-based MQANOVA offers competitive power across distributions and correlation levels: while MANOVA is the strongest performer under normality, its power

degrades substantially under non-normal conditions where MQANOVA-Permutation maintains a clear, and sometimes large, advantage.

v. Anthropometric profiles of Pakistani children and adolescents are strongly stratified by geography, with Lahore showing consistently elevated height, weight, and circumference measures relative to Multan and the Rawalpindi/Islamabad corridor, confirmed by a highly significant Location effect ($p < 0.0001$) across the entire quantile spectrum.

vi. Gender-based growth disparities are moderated by location: the Location × Gender interaction ($p = 0.0022$) shows the weight gap ranging from 0.5 kg in Lahore to 1.3 kg in Multan, consistent with greater nutritional disadvantage for girls in less-resourced settings.

vii. The close agreement between asymptotic and permutation p-values for symmetric, large-sample effects (e.g., the Location effect) validates the mathematical stability of the MQANOVA test statistic, while the divergence for skewed traits such as waist circumference ($p < 0.0001$ asymptotic vs. $p = 0.0202$ permutation) demonstrates the practical value of permutation-based inference in guarding against false-positive findings.

4.3 RECOMMENDATIONS

On methodological grounds, the permutation-based MQANOVA is recommended as the default procedure for multivariate factorial designs whenever outcomes may be non-normal, heteroscedastic, or the design unbalanced, given its consistently superior Type I error control and competitive power relative to asymptotic MQANOVA, classical MANOVA, and Separate QANOVA. Classical MANOVA remains an appropriate and efficient choice only when data are approximately multivariate normal and homoscedastic, conditions that should be checked rather than assumed. Among covariance-estimation strategies, a kernel-density-based estimator is recommended as the default owing to its favorable stability and power relative to purely interval-based alternatives.

On public health grounds, the finding that the gender gap in weight is nearly three times larger in Multan (1.3 kg) than in Lahore (0.5 kg) supports region-specific, rather than uniform national, nutrition and growth-monitoring policies, with priority given to interventions addressing the nutritional disadvantage of girls in Multan and to obesity-prevention strategies appropriate to Lahore's upper-tail weight profile. On software-development grounds, incorporating the MQANOVA methodology into widely used statistical packages with accessible functions, parallelized permutation computation, and built-in visualization would materially lower the barrier to adoption by applied researchers.

For future methodological work, promising directions include adaptive quantile selection to target distributional regions of specific applied interest; extension to longitudinal and repeated-measures designs; regularization strategies for higher-dimensional outcome vectors ($d > 5$); generalization to mixed continuous and categorical outcomes; small-sample and Bayesian approaches to missing data; extensions to causal and quantile treatment-effect estimation in observational studies; and computational optimization, including GPU-accelerated permutation routines and comparison against machine-learning-based alternatives. Such extensions would broaden the applicability of the MQANOVA framework while preserving its core strengths of robustness, interpretability, and valid inference under realistic, imperfect data conditions.

REFERENCES

Anderson, T. W. (2003). An Introduction to Multivariate Statistical Analysis (3rd ed.). Wiley.

Bahadur, R. R. (1966). A note on quantiles in large samples. Annals of Mathematical Statistics, 37(3), 577–580.

Brunner, E., Bathke, A. C., & Konietschke, F. (2018). Rank and Pseudo-Rank Procedures for Independent Observations in Factorial Designs. Springer.

Ditzhaus, M., Fried, R., & Pauly, M. (2021). QANOVA: Quantile-Based Permutation Methods for General Factorial Designs. TEST, 30(4), 960–979.

Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics (4th ed.). Sage.

Fisher, R. A. (1925). Statistical Methods for Research Workers. Oliver & Boyd.

Friedrich, S., Konietschke, F., & Pauly, M. (2017). Resampling-Based Analysis of Factorial Designs. Journal of Statistical Planning and Inference, 189, 59–72.

Hallin, M., Paindaveine, D., & Šiman, M. (2010). Multivariate Quantiles and Multiple-Output Regression Quantiles. Bernoulli, 16(1), 143–169.

Hettmansperger, T. P., & McKean, J. W. (2011). Robust Nonparametric Statistical Methods (2nd ed.). CRC Press.

Huber, P. J. (1981). Robust Statistics. Wiley.

Janssen, A. (1997). Studentized Permutation Tests for Non-IID Hypotheses. Statistics & Probability Letters, 36, 9–21.

Johnson, R. A., & Wichern, D. W. (2019). Applied Multivariate Statistical Analysis (6th ed.). Pearson.

Koenker, R. (2005). Quantile Regression. Cambridge University Press.

Koenker, R., & Bassett, G. (1978). Regression Quantiles. Econometrica, 46(1), 33–50.

Koenker, R., & Hallock, K. F. (2001). Quantile Regression. Journal of Economic Perspectives, 15(4), 143–156.

Kong, L., & Mizera, I. (2012). Quantile Tomography. Annals of Statistics, 40(6), 2998–3020.

Montgomery, D. C. (2019). Design and Analysis of Experiments (10th ed.). Wiley.

Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. Journal of the American Statistical Association, 69(348), 894–908.

Pauly, M., Brunner, E., & Konietschke, F. (2015). Asymptotic Permutation Tests in General Factorial Designs. Journal of the Royal Statistical Society Series B, 77(2), 461–473.

Powell, J. L. (1986). Censored regression quantiles. Journal of Econometrics, 32(1), 143–155.

R Core Team. (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Rencher, A. C., & Christensen, W. F. (2012). Methods of Multivariate Analysis (3rd ed.). Wiley.

Serfling, R. J. (1980). Approximation Theorems of Mathematical Statistics. Wiley.

Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis. Chapman & Hall.

Wilcox, R. R. (2012). Introduction to Robust Estimation and Hypothesis Testing (3rd ed.). Academic Press.

Wilcox, R. R. (2017). Modern Statistics for the Social and Behavioral Sciences. CRC Press.