



MerQur'da Parametrik Testler: Tek Örneklem t-Testinden MANOVA'ya Tüm Aile

Parametric Tests in MerQur: The Full Family from One-Sample t-Test to MANOVA

Ömer K. Örucü¹

¹ Süleyman Demirel Üniversitesi, Mimarlık Fakültesi, Peyzaj Mimarlığı Bölümü, Isparta/Türkiye ORCID:

0000-0002-2162-7553 E-posta: omerorucu@sdu.edu.tr · Resmi site:

<https://www.sekizgenacademy.com/journals/index.php/merqur/tr/index>

Yazışmadan sorumlu yazar: Ömer K. Örucü (omerorucu@sdu.edu.tr)

Tür: Davetli Editöryal Sunum / Invited Editorial Showcase **Geliş:** 2026-05-17 · **Kabul:** 2026-05-17 · **Yayın:** 2026-05-17 **DOI:** — (ISSN başvurusu sonrası eklenecek)

Öz

Parametrik testler, gözlemlerin belirli bir parametrik dağılımdan (genellikle normal) geldiği varsayımı altında, popülasyon parametrelerine ilişkin hipotezleri test eden yöntemler ailesidir. Akademik araştırmaların büyük çoğunluğunda hâlâ en sık kullanılan analiz tipi olan parametrik testler, doğru uygulandığında en yüksek istatistiksel güce sahip yöntemleri sağlar. Bu çalışmada, **MerQur** masaüstü yazılımının Parametrik Testler kategorisinde sunduğu **11 analiz** ayrıntılı olarak tanıtılmıştır: tek örneklem t-testi, bağımsız iki örneklem t-testi, eşleştirilmiş t-testi, tek yönlü ANOVA, iki yönlü ANOVA, tekrarlı ölçümler ANOVA, MANOVA, ANCOVA, bootstrap güven aralığı, permütasyon testi ve çoklu karşılaştırma düzeltme. Her analiz için (i) test edilen hipotez ve uygulama bağlamı, (ii) gerekli varsayımlar (normallik, varyans homojenliği, bağımsızlık, küresellik), (iii) MerQur'daki form alanları ve parametre seçenekleri, (iv) raporlanan istatistikler ve etki büyüklükleri, ve (v) tipik bir araştırma sorusu için yorumlama önerisi sunulmuştur. Bootstrap CI ve permütasyon testi, klasik parametrik testlerin varsayımları sağlanmadığında yeniden örneklemeli alternatif sunması nedeniyle aynı kategoride yer almıştır. Çoklu karşılaştırma düzeltme bölümünde Bonferroni, Holm, Hochberg, Benjamini-Hochberg (FDR), Sidak gibi yöntemler ve uygun kullanım koşulları tartışılmıştır. Sonuç olarak MerQur'un Parametrik Testler kategorisi, basit grup karşılaştırmasından çok değişkenli kovaryanslı tasarımlara uzanan geniş bir araştırma yelpazesini, varsayım kontrolleri ve etki büyüklükleri ile birlikte tek bir grafik arayüzde sunmaktadır.

Anahtar Kelimeler: parametrik testler, t-testi, ANOVA, MANOVA, ANCOVA, bootstrap, permütasyon testi, çoklu karşılaştırma, MerQur

Abstract

Parametric tests are a family of methods that test hypotheses about population parameters under the assumption that observations follow a parametric distribution (typically normal). Still the most frequently used analysis type in academic research, parametric tests — when applied correctly — provide the methods with the highest statistical power. This study introduces in detail the **11 analyses** offered under the Parametric Tests category of **MerQur**: one-sample t-test, independent two-sample t-test, paired t-test, one-way ANOVA, two-way ANOVA, repeated-measures ANOVA, MANOVA, ANCOVA, bootstrap confidence interval, permutation test, and multiple-comparison correction. For each analysis, the following are presented: (i) the hypothesis tested and application context, (ii) required assumptions (normality, homogeneity of variance, independence, sphericity), (iii) form fields and parameter options in MerQur, (iv) reported statistics and effect sizes, and (v) an interpretation guide for a typical research question. Bootstrap CI and permutation tests are included in the same category as they offer resampling-based alternatives when assumptions of classical parametric tests are violated. The multiple-comparison correction section discusses Bonferroni, Holm, Hochberg, Benjamini-Hochberg (FDR), and Sidak methods together with appropriate-use criteria. Overall, MerQur's Parametric Tests category presents a wide research spectrum — from simple group comparison to multivariate designs with covariates — together with assumption checks and effect sizes within a single graphical interface.

Keywords: parametric tests, t-test, ANOVA, MANOVA, ANCOVA, bootstrap, permutation test, multiple comparison, MerQur

1. Giriş

Parametrik testler, çıkarımsal istatistiğin köşe taşıdır. Yöntemler, gözlenen veriden popülasyon parametrelerine (ortalama, varyans, kovaryans) ilişkin olasılıksal çıkarım yapmamızı sağlar — koşul: verinin parametrik bir dağılımdan (genellikle normal veya yaklaşık-normal) geldiği varsayımı geçerli olmalıdır. Bu varsayım sağlandığında, parametrik testler aynı veri kümesi üzerinde parametrik olmayan rakiplerinden **daha yüksek istatistiksel güce** sahip olur (Cohen, 1988); yani aynı etkiyi daha küçük örneklem büyüklüğüyle tespit edebilirler. Bu nedenle, varsayımları sağlanan veriler için parametrik testler tercih edilmesi gereken yöntemlerdir.

Ne yazık ki literatürde p-değeri ve istatistiksel anlamlılığın yanlış raporlanması; varsayım kontrollerinin atlanması; çoklu karşılaştırmanın yapılmaması veya yanlış düzeltme uygulanması yaygın sorunlardır (Nuijten ve ark., 2016). Bir yazılımın değeri, doğru analizi seçmenin ötesine geçerek, varsayım kontrollerini görsel olarak sunması ve düzeltme yöntemlerini kullanıcıya hatırlatması ile de ölçülür. Bu çalışmanın amacı, **MerQur** masaüstü yazılımının Parametrik Testler kategorisinde sunduğu 11 analizi ayrıntılı olarak tanıtmak; her birinin hangi araştırma sorusuna cevap verdiğini, hangi varsayımları gerektirdiğini ve MerQur arayüzünde nasıl uygulandığını göstermektir.

2. Tek Örneklem t-Testi

2.1 Yöntem

Tek örneklem t-testi, bir örneklem ortalamasının önceden belirlenmiş bir referans değere (μ_0) eşit olup olmadığını test eder. Test istatistiği:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}, \quad df = n - 1$$

burada \bar{x} örneklem ortalaması, s örneklem standart sapması, n örneklem büyüklüğüdür. Sıfır hipotezi

$H_0: \mu = \mu_0$, alternatif $H_1: \mu \neq \mu_0$ (iki yönlü) veya $H_1: \mu > \mu_0 / \mu < \mu_0$ (tek yönlü) olabilir.

2.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Tek Örneklem t-Testi

. Form alanları:

- **Değer Sütunu:** test edilecek sayısal değişken
- **Test edilen μ :** referans değer (sayı)
- **Alternatif hipotez:** İki yönlü / $X > \mu$ / $X < \mu$
- **Güven düzeyi:** %95 (varsayılan), %90, %99

Çıktıda t istatistiği, serbestlik derecesi, p-değeri, ortalama farkı, %95 güven aralığı ve **Cohen's d** etki büyüklüğü raporlanır. Varsayım kontrolü olarak Shapiro-Wilk normallik testi otomatik olarak çalıştırılır; normallikten anlamlı sapma varsa kullanıcı uyarılır ve alternatif olarak Wilcoxon işaretli sıra testi önerilir.

2.3 Uygulama örneği

Bir orman fidanlığında yetiştirilen 30 çam fidanının ortalama boyunun, bölgesel hedef olan 25 cm'den farklı olup olmadığı test edilebilir. Sonuç: $t(29) = -2.15$, $p = 0.040$, ortalama fark = -1.8 cm, %95 GA = [-3.5, -0.1], Cohen's d = -0.39. Yorum: hedef değerden istatistiksel olarak anlamlı düşüş; etki büyüklüğü küçük-orta.

3. Bağımsız İki Örneklem t-Testi

3.1 Yöntem

Bağımsız iki örneklem t-testi, iki farklı grubun ortalamalarının eşit olup olmadığını test eder. **Student t-testi** (eşit varyans varsayımı) ve **Welch t-testi** (varyansların eşitsiz olabileceği genelleme) iki ana versiyonudur. Welch t-testi günümüzde varsayılan olarak önerilmektedir çünkü varyans homojenliği sağlandığında Student'a yakın güç, sağlanmadığında daha doğru tip-1 hata kontrolü sunar (Delacre ve ark., 2017).

3.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Bağımsız t-Testi

. Form alanları:

- **Grup Sütunu:** iki kategorili grup değişkeni
- **Değer Sütunu:** karşılaştırılacak sayısal değişken
- **Varyans:** Eşit (Student) / Eşit değil (Welch) — varsayılan Welch
- **Alternatif hipotez:** iki yönlü/tek yönlü

Çıktıda her grup için tanımlayıcı istatistikler (n , ortalama, SD, SE), t istatistiği, serbestlik derecesi, p-değeri, ortalama farkı, %95 GA ve **Cohen's d + Hedges' g** raporlanır. Varsayım kontrolü olarak Shapiro-Wilk (her grup için ayrı) ve Levene homojenlik testi otomatik çalıştırılır.

3.3 Uygulama örneği

Eğitilmiş ve kontrol grubu öğrencilerinin bir ön-test puanları karşılaştırılır. $n_1 = 40$, $n_2 = 38$; Welch $t(75.2) = 3.18$, $p = 0.002$; ortalama fark = 4.5 puan; Cohen's $d = 0.72$ (orta-büyük etki).

4. Eşleştirilmiş t-Testi

4.1 Yöntem

Eşleştirilmiş (paired) t-testi, aynı deneklerden farklı zaman noktalarında veya farklı koşullarda alınan iki ölçümün ortalamalarını karşılaştırır. Aslında **fark sütunu** üzerinden tek örneklem t-testi olarak yürütülür ($H_0: \mu_d = 0$). Eşleştirme, denek-içi varyansın kontrol edilmesini sağlayarak istatistiksel gücü artırır.

4.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Eşleştirilmiş t-Testi

. Form alanları:

- **1. Ölçüm:** ilk ölçüm sütunu
- **2. Ölçüm:** ikinci ölçüm sütunu (aynı satır = aynı denek)
- **Alternatif hipotez:** iki yönlü / artış / azalış

Çıktıda fark için ortalama, SD, t istatistiği, df, p-değeri, %95 GA ve **Cohen's dz** raporlanır. Fark sütununun normalliği Shapiro-Wilk ile test edilir.

4.3 Uygulama örneği

Bir uygulamadan önce ve sonra ölçülen ağrı skorları ($n=22$ hasta). Ortalama fark = -2.7 (sonra düşük), $t(21) = -4.85$, $p < 0.001$, Cohen's $d = -1.03$ (büyük etki). Tedavi sonrası ağrı anlamlı azalmıştır.

5. Tek Yönlü ANOVA

5.1 Yöntem

Tek yönlü ANOVA (One-Way Analysis of Variance), üç veya daha fazla bağımsız grup ortalamasının eşit olup olmadığını test eder. F istatistiği, gruplar arası varyans ile grup içi varyansın oranıdır:

$$F = \frac{MS_{\text{gruplar arası}}}{MS_{\text{grup içi}}}, \quad df_1 = k - 1, \quad df_2 = N - k$$

Anlamlı F bulgusu, gruplar arası en az bir farkın istatistiksel olarak anlamlı olduğunu gösterir; ancak hangi grupların farklı olduğunu söylemez. Bu nedenle **post-hoc testler** (Tukey HSD, Scheffé, Bonferroni, Games-Howell) gereklidir.

5.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Tek Yönlü ANOVA

. Form alanları:

- **Grup Sütunu:** kategorik grup değişkeni (3+ seviye)

- **Değer Sütunu:** sayısal yanıt değişkeni
- **Post-hoc:** Tukey HSD / Scheffé / Bonferroni / Games-Howell (varyans homojen değilse)
- **Etki büyüklüğü:** η^2 , ω^2 , parsiyel η^2

Çıktıda ANOVA tablosu (SS, df, MS, F, p), etki büyüklükleri, Levene testi, post-hoc karşılaştırma matrisi ve **gruplar arası kutu grafiği** sunulur.

5.3 Uygulama örneği

Üç farklı gübre dozajının (düşük, orta, yüksek) bitki büyümesi üzerindeki etkisi. $n_{per_group} = 20$. $F(2, 57) = 8.42$, $p < 0.001$, $\eta^2 = 0.23$ (büyük etki). Tukey HSD: yüksek > düşük ($p < 0.001$), yüksek > orta ($p = 0.018$), orta vs düşük ($p = 0.087$ — anlamsız).

6. İki Yönlü ANOVA

6.1 Yöntem

İki yönlü ANOVA, iki kategorik faktörün (A ve B) bir sayısal yanıt değişkeni üzerindeki ana etkilerini ve **etkileşim** (A×B) etkisini eş zamanlı olarak test eder. Etkileşim, bir faktörün etkisinin diğer faktörün seviyesine bağlı olarak değişip değişmediğini gösterir.

6.2 MerQur'da uygulama

İstatistik → Parametrik Testler → İki Yönlü ANOVA

. Form alanları:

- **Faktör 1:** birinci kategorik faktör
- **Faktör 2:** ikinci kategorik faktör
- **Bağımlı Değişken:** sayısal yanıt
- **Etkileşim:** dahil / hariç (varsayılan dahil)
- **Tip:** Tip I, II, III sums of squares (varsayılan Tip III)

Çıktıda ANOVA tablosu (her ana etki + etkileşim için F, p, η^2), **etkileşim grafiği** (mean \pm SE), Levene testi sunulur. Etkileşim anlamlı çıkarsa, basit ana etkiler (simple main effects) ek olarak hesaplanır.

6.3 Uygulama örneği

İki tedavi yöntemi (T1, T2) × iki cinsiyet (E, K) faktöriyel tasarımı; bağımlı değişken sağlık skoru. Tedavi ana etkisi anlamlı, cinsiyet ana etkisi anlamsız, tedavi×cinsiyet etkileşimi anlamlı ($F=5.3$, $p=0.024$) — yorum: tedavi etkisi cinsiyete göre farklılaşıyor.

7. Tekrarlı Ölçümler ANOVA

7.1 Yöntem

Tekrarlı ölçümler ANOVA (Repeated Measures ANOVA, RM-ANOVA), aynı deneklerden 3+ koşulda veya zaman noktasında alınan ölçümleri karşılaştırır. Denek-içi varyansı kontrol etmesi sebebiyle bağımsız ANOVA'dan

daha güçlüdür. Önemli ek varsayım **küresellik** (sphericity): tüm koşul çiftleri arasındaki fark varyanslarının eşit olması. Mauchly küresellik testi ihlali olursa, serbestlik dereceleri **Greenhouse-Geisser** veya **Huynh-Feldt** düzeltmesiyle azaltılır.

7.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Tekrarlı Ölçümler ANOVA

. Form alanları:

- **Ölçüm Sütunları:** her satırın aynı deneğe ait olduğu, koşul başına bir sütun
- **Küresellik düzeltmesi:** Yok / Greenhouse-Geisser / Huynh-Feldt (Mauchly anlamlıysa otomatik öneri)
- **Post-hoc:** Bonferroni / Holm

Çıktıda Mauchly testi, ANOVA tablosu (gerekirse düzeltilmiş df ile), parsiyel η^2 , post-hoc çift karşılaştırmaları ve **zaman grafiği** sunulur.

7.3 Uygulama örneği

Bir aylık egzersiz programı sırasında 4 farklı zamanda ölçülen kalp atış hızı dinlenme değerleri ($n = 25$ katılımcı). $F(3, 72) = 12.4$, $p < 0.001$, parsiyel $\eta^2 = 0.34$. Bonferroni post-hoc: zaman 4 < zaman 1 ($p < 0.001$).

8. MANOVA

8.1 Yöntem

MANOVA (Multivariate Analysis of Variance), birden çok bağımlı değişkeni eş zamanlı olarak bir veya daha fazla kategorik faktöre karşı test eder. Tek değişkenli ANOVA'ları ayrı ayrı çalıştırmaya göre üstünlüğü: (i) tip-1 hatayı kontrol eder, (ii) bağımlı değişkenler arası korelasyonu modele dahil eder. Test istatistikleri: **Wilks's A** (en yaygın), Pillai's trace, Hotelling-Lawley, Roy's largest root.

8.2 MerQur'da uygulama

İstatistik → Parametrik Testler → MANOVA

. Form alanları:

- **Grup (Faktör):** kategorik grup değişkeni
- **Bağımlılar:** birden çok sayısal yanıt değişkeni
- **Test istatistiği:** Wilks / Pillai / Hotelling / Roy (varsayılan Wilks)

Çıktıda MANOVA tablosu (her istatistik için yaklaşık F, p), her bağımlı değişken için ayrı ANOVA çıktısı (takip analizi olarak) ve **çok değişkenli etki büyüklüğü** (parsiyel η^2) sunulur. Varsayım kontrolleri: çok değişkenli normallik (Mardia testi), kovaryans matrisi homojenliği (Box's M).

8.3 Uygulama örneği

Üç eğitim yönteminin (geleneksel, hibrit, çevrimiçi) hem matematik hem sözel performans üzerindeki etkisi. Wilks's $\Lambda = 0.74$, $F(4, 152) = 6.18$, $p < 0.001$, parsiyel $\eta^2 = 0.14$. Takip ANOVA: matematik anlamlı, sözel anlamsız.

9. ANCOVA

9.1 Yöntem

ANCOVA (Analysis of Covariance), kategorik bir faktörün etkisini test ederken bir veya daha fazla **sürekli kovaryansı** istatistiksel olarak kontrol eder. Kovaryans, yanıt değişkeni ile ilişkili ama deneysel ilgi alanı dışındaki bir değişkendir (ör. ön-test puanı, yaş, IQ). Kovaryansı modele dahil etmek, hata varyansını azaltır ve gerçek grup etkisinin tahminini iyileştirir. Önemli varsayım: **regresyon eğimlerinin homojenliği** — kovaryans ile yanıt arasındaki ilişki tüm gruplarda benzer olmalı.

9.2 MerQur'da uygulama

İstatistik → Parametrik Testler → ANCOVA

. Form alanları:

- **Bağımlı değişken (sürekli)**: sayısal yanıt
- **Faktör (kategorik)**: grup değişkeni
- **Kovaryans (sürekli, ≥ 1)**: bir veya daha fazla sürekli kovaryans
- **Düzeltilmiş ortalamalar**: dahil/hariç (varsayılan dahil)

Çıktıda ANCOVA tablosu, **düzeltilmiş (adjusted) ortalamalar** (kovaryansın ortalama değerinde tahmin edilen), parsiyel η^2 , regresyon eğimi homojenliği testi (faktör \times kovaryans etkileşimi) sunulur.

9.3 Uygulama örneği

İki müdahalenin son-test puanlarına etkisi; ön-test puanı kovaryans. Düzeltilmiş ortalamalar: müdahale A = 78.2, müdahale B = 74.6; $F(1,56) = 4.8$, $p = 0.033$, parsiyel $\eta^2 = 0.08$. Ön-test kontrolünden sonra A grubu anlamlı üstün.

10. Bootstrap Güven Aralığı

10.1 Yöntem

Bootstrap, parametrik dağılım varsayımı yapmadan örneklem dağılımını **veri üzerinden yeniden örnekleme** ile tahmin eder (Efron, 1979). Klasik t-tabanlı güven aralıkları normallik varsayımına dayanırken, bootstrap CI bu varsayıma ihtiyaç duymaz; özellikle küçük veya çarpık dağılımlı örneklemler için güvenilir alternatiftir. **Percentile bootstrap**, **bias-corrected and accelerated (BCa)** ve **basic bootstrap** olmak üzere üç ana CI yöntemi mevcuttur.

10.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Bootstrap CI

. Form alanları:

- **Değişken(ler):** istatistiğin hesaplanacağı sütun(lar)
- **İstatistik:** ortalama / medyan / SD / korelasyon / ortalama farkı
- **CI tipi:** Percentile / BCa / Basic
- **Bootstrap replikat sayısı:** 1000 (varsayılan), 2000, 5000, 10000
- **Güven düzeyi:** %95 (varsayılan), %90, %99

Çıktıda örneklem istatistiği, bootstrap dağılımının histogramı, ortalama, SD, %95 CI ve bias tahmini sunulur.

10.3 Uygulama örneği

Çarpık dağılımlı gelir verisinin medyanı için %95 CI. Örneklem medyan = 32.500 TL, 5000 bootstrap BCa CI = [29.800; 36.100]. Klasik parametrik formülün uygulanamayacağı durumlarda güvenilir tahmin.

11. Permütasyon Testi

11.1 Yöntem

Permütasyon (randomization) testi, sıfır hipotezi altında etiketlerin değiştirilebilir olduğunu varsayar (exchangeability) ve gözlenen test istatistiğinin tüm olası etiket yeniden düzenlemeleri altındaki dağılımına karşı değerini değerlendirir. Klasik testler gibi normallik veya varyans homojenliği gerektirmez; örneklem büyüklüğü küçük olduğunda veya dağılım bilinmediğinde güvenilir bir alternatif sunar.

11.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Permutation Test

. Form alanları:

- **Grup Sütunu:** karşılaştırma grubu
- **Değer Sütunu:** sayısal değişken
- **İstatistik:** ortalama farkı / medyan farkı / t istatistiği
- **Permütasyon sayısı:** 1000 (varsayılan), 5000, 10000, exact (n küçükse)
- **Alternatif hipotez:** iki yönlü / tek yönlü

Çıktıda gözlenen istatistik, permütasyon dağılımının histogramı (gözlenen değer kırmızı çizgi ile), permütasyon p-değeri ve %95 permütasyon CI sunulur.

11.3 Uygulama örneği

Küçük örneklemlili, çarpık dağılımlı verilerin iki grup karşılaştırması. $n_1 = 8$, $n_2 = 7$; klasik t-test'in normallik varsayımı sağlanmıyor. Permütasyon $p = 0.018$ (10 000 permütasyon) — anlamlı fark.

12. Çoklu Karşılaştırma Düzeltme

12.1 Yöntem

Aynı veri kümesi üzerinde birden çok test çalıştırıldığında, en az bir tip-1 hata yapma olasılığı (family-wise error rate, FWER) bireysel test α düzeyini aşar. Bu sorunu yönetmek için **çoklu karşılaştırma düzeltme** yöntemleri uygulanır:

- **Bonferroni:** en katı, en muhafazakâr; α 'yı test sayısına böler
- **Holm:** Bonferroni'nin sıralı (step-down) versiyonu; daha güçlü, FWER kontrolü
- **Hochberg:** sıralı (step-up); pozitif bağımlı testlerde daha güçlü
- **Sidak:** bağımsız testler için Bonferroni'den biraz daha güçlü
- **Benjamini-Hochberg (FDR):** yanlış keşif oranını kontrol eder, çok sayıda hipotez (genomik vb.) için tercih

12.2 MerQur'da uygulama

İstatistik → Parametrik Testler → Çoklu Karşılaştırma Düzeltme

. Form alanları:

- **p-değerleri sütunu** (veya manuel giriş)
- **Yöntem:** Bonferroni / Holm / Hochberg / Sidak / BH (FDR) / BY (FDR — bağımlı)
- **α düzeyi:** 0.05 (varsayılan)

Çıktıda her hipotez için ham ve düzeltilmiş p-değerleri, anlamlılık kararı (red/kabul) ve grafiksel “Manhattan benzeri” sıralı p-değer grafiği sunulur.

12.3 Uygulama örneği

20 farklı endpoint üzerinde test yapılmış bir klinik çalışma. Ham 6 p-değeri anlamlı görünüyor; BH-FDR düzeltmesi sonrası 3 anlamlı kalıyor.

13. Karşılaştırmalı Değerlendirme ve Sınırlılıklar

Tablo 1, MerQur'un Parametrik Testler kategorisindeki 11 analiz için diğer GUI tabanlı açık erişimli alternatiflerle karşılaştırmasını özetlemektedir.

Tablo 1. Parametrik test ailesinin MerQur, JASP, jamovi ve PSPP'deki desteği.

Analiz	MerQur	JASP	jamovi	PSPP
Tek Örneklem t	✓	✓	✓	✓
Bağımsız t (Welch)	✓	✓	✓	✓
Eşleştirilmiş t	✓	✓	✓	✓
Tek Yönlü ANOVA	✓	✓	✓	✓
İki Yönlü ANOVA	✓	✓	✓	+
Tekrarlı Ölçümler ANOVA	✓	✓	✓	-
MANOVA	✓	+	+	-
ANCOVA	✓	✓	✓	-
Bootstrap CI	✓	+	+	-
Permütasyon Testi	✓	-	-	-
Çoklu Karşılaştırma Düzeltme	✓	+	+	-

MerQur'un belirgin avantajı: **MANOVA, permütasyon testi ve çoklu karşılaştırma düzeltmesinin** tam destekli olarak sunulması, klasik t/ANOVA ailesinin tüm üyelerinin tek bir kategoride toplanması.

Sınırlılıklar: (i) Çok faktörlü ANOVA (3+ faktör) için doğrudan panel bulunmaz — bu tür tasarımlar için Karma Modeller (LMM) kategorisi kullanılır; (ii) Latin kare, split-plot ve RCBD gibi özel deneysel tasarımlar için doğrudan panel yoktur — bunların VARCOMP veya LMM ile modellenmesi önerilir; (iii) Welch ANOVA (varyans homojenliği sağlanmadığında) doğrudan panel olarak yer almaz, ancak post-hoc'ta Games-Howell seçeneği bu durumu kısmen karşılar.

14. Sonuç

Bu çalışma, MerQur masaüstü yazılımının Parametrik Testler kategorisinde sunulan 11 analizi sistemli olarak tanıtmıştır. Tek örneklem t-testinden MANOVA'ya, klasik tasarımlardan bootstrap/permütasyon yeniden örnekleme yöntemlerine, çoklu karşılaştırma düzeltmesine uzanan kapsam; akademik araştırmalarda en sık başvurulan analiz ailesini tek bir grafik arayüzde sunar. Her analizde varsayım kontrolleri otomatik çalışır; uygun olduğunda etki büyüklükleri ve güven aralıkları raporlanır; post-hoc karşılaştırmalar APA 7'ye yakın formatta sunulur. MerQur'un Parametrik Testler kategorisi, lisans öğrencilerinden deneyimli araştırmacılara kadar geniş bir kullanıcı kitlesinin doğru yöntemi seçip uygulamasına olanak veren bir referans noktasıdır. Sonraki davetli editöryal sunumlarda non-parametrik testler, ilişki, regresyon, sınıflandırma, kümeleme, ileri düzey ve sağkalım kategorileri benzer ayrıntıyla incelenecektir.

Beyanlar

Etik Kurul Onayı: Bu çalışma insan ya da hayvan denek içermediğinden etik kurul onayı gerektirmemiştir.

Çıkar Çatışması: Yazar, MerQur yazılımının geliştiricisidir. Bu durum dışında herhangi bir çıkar çatışması bildirmemiştir.

Finansman: Spesifik bir dış fon alınmamıştır.

Veri ve Kod Erişim Beyanı: Bu derleme orijinal araştırma verisi içermez. MerQur yazılımı

<https://merqur.sdu.edu.tr> adresinden ücretsiz indirilebilir.

Yapay Zekâ Kullanımı: Bu makalenin yazımı sırasında üretken yapay zekâ (Claude, Anthropic) dil ve yapı düzeltmesi amacıyla destekleyici olarak kullanılmıştır. Bilimsel içerik yazarın özgün katkısıdır.

Yazar Katkı Beyanı (CRediT): Ömer K. Örüçü — Kavramsallaştırma, Yöntem, Yazılım, Doğrulama, Yazma (orijinal taslak), Yazma (gözden geçirme & düzenleme).

Kaynakça

Cohen, J. (1988).

Statistical power analysis for the behavioral sciences
(2nd ed.). Lawrence Erlbaum.

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test.

International Review of Social Psychology
, 30(1), 92–101. <https://doi.org/10.5334/irsp.82>

Efron, B. (1979). Bootstrap methods: Another look at the jackknife.

Annals of Statistics
, 7(1), 1–26. <https://doi.org/10.1214/aos/1176344552>

Efron, B., & Tibshirani, R. J. (1993).

An introduction to the bootstrap
. Chapman & Hall.

Fisher, R. A. (1925).

Statistical methods for research workers
. Oliver & Boyd.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data.

Psychometrika
, 24(2), 95–112. <https://doi.org/10.1007/BF02289823>

Hedges, L. V., & Olkin, I. (1985).

Statistical methods for meta-analysis
. Academic Press.

Holm, S. (1979). A simple sequentially rejective multiple test procedure.

Scandinavian Journal of Statistics

, 6(2), 65–70.

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs.

Journal of Educational Statistics

, 1(1), 69–82. <https://doi.org/10.3102/10769986001001069>

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs.

Frontiers in Psychology

, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution.

Annals of Mathematical Statistics

, 11(2), 204–209. <https://doi.org/10.1214/aoms/1177731915>

Nuijten, M. B., Hartgerink, C. H. J., van Assen, M. A. L. M., Epskamp, S., & Wicherts, J. M. (2016). The prevalence of statistical reporting errors in psychology (1985–2013).

Behavior Research Methods

, 48(4), 1205–1226. <https://doi.org/10.3758/s13428-015-0664-2>

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs.

Psychological Methods

, 8(4), 434–447. <https://doi.org/10.1037/1082-989X.8.4.434>

Pernet, C. R., Wilcox, R., & Rousseelet, G. A. (2013). Robust correlation analyses: False positive and power validation using a new open source matlab toolbox.

Frontiers in Psychology

, 3, 606. <https://doi.org/10.3389/fpsyg.2012.00606>

Phipson, B., & Smyth, G. K. (2010). Permutation P-values should never be zero: Calculating exact P-values when permutations are randomly drawn.

Statistical Applications in Genetics and Molecular Biology

, 9(1), 39. <https://doi.org/10.2202/1544-6115.1585>

Pituch, K. A., & Stevens, J. P. (2016).

Applied multivariate statistics for the social sciences

(6th ed.). Routledge.

Rice, W. R. (1989). Analyzing tables of statistical tests.

Evolution

, 43(1), 223–225. <https://doi.org/10.1111/j.1558-5646.1989.tb04220.x>

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions.

Journal of the American Statistical Association

, 62(318), 626–633. <https://doi.org/10.2307/2283989>

Tabachnick, B. G., & Fidell, L. S. (2019).

Using multivariate statistics

(7th ed.). Pearson.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance.

Biometrics

, 5(2), 99–114. <https://doi.org/10.2307/3001913>

Welch, B. L. (1947). The generalization of “Student’s” problem when several different population variances are involved.

Biometrika

, 34(1–2), 28–35. <https://doi.org/10.1093/biomet/34.1-2.28>

Bu makale “Davetli Editöryal Sunum” bölümü kapsamında yayımlanmıştır. Bölüm politikası gereği harici hakem değerlendirmesinden geçmemiş, MerQur Veri Bilimi ve Yöntemleri Dergisi Yayın Kurulu tarafından editöryal incelemeye tabi tutulmuştur. Bu makale Creative Commons Atıf 4.0 Uluslararası (CC-BY 4.0) lisansı altında yayımlanmıştır.