

HİPOTEZ (ÖNEMLİLİK) TESTLERİ

Test seçimi ve
yorumlama

Doç.Dr.Hasibe Kadiođlu

AMAÇ

- Önemlilik testlerinin kullanıldığı durumların ve test sonucunun yorumlanmasının öğrenilmesi

HEDEFLER

- Parametrik ve non-parametrik testleri sıralar.
- Bağımlı ve bağımsız grup arasındaki farkı açıklar.
- Parametrik test varsayımlarını sıralar.
- Analiz için gerekli olan önemlilik testini seçer.
- Önemlilik testinin sonucunu yorumlar.
- Testin sonucunda karar cümlesini yazar.

İÇERİK

- Önemlilik testlerinin sınıflandırması.
- Parametrik test varsayımları
- Önemlilik test seçimi basamakları
- Önemlilik test sonuçlarının yorumlanması ve karar cümlesinin yazılması.

HİPOTEZ TESTLERİ

```
graph TD; A[HİPOTEZ TESTLERİ] --> B[Test seçimi]; A --> C[Test sonucunu yorumlama];
```

Test seçimi

Test
sonucunu
yorumlama

ÖNEMLİLİK TESTLERİ

| Parametrik testler | Non parametrik testler | |
|--|------------------------|-------------------------|
| Kitle ortalamasının önemlilik testi (One –Sample T test) | İşaret testi | Ki-kare (Chi-Square) |
| İki ortalama arasındaki farkın önemlilik test (Independent Sample T test) | Mann-Whitney U testi | Fisher’s Exact test |
| Tek yönlü varyans (One-way ANOVA) | Kruskal-wallis testi | Çok gözlü Ki-kare |
| İki eş arasındaki farkın önemlilik testi (Paired- Sample T test) | Wilcoxon testi | Mc Nemar testi |
| Tekrarlı ölçümlerde varyans analizi Repeated-measures ANOVA | Friedman test | Cochran Q testi |

NEYE GÖRE
HANGİ TESTİ SEÇECEĞİZ?

Önemlilik testi;

1. Verinin özelliğine (parametrik?, nonparametrik?),
2. Örnek Büyüklüğüne (gruplardaki denek sayısı),
3. İncelenen grupların bağımlı ya da bağımsız olmasına,
4. Karşılaştırılacak grup sayısına,
5. Verinin normal dağılıma uygun olup olmamasına göre seçilir

| Gruplar | | Veri Tipi | | |
|---------------|------------------------|---|--|--|
| | | Parametrik veri | | Non parametrik veri |
| | | Parametrik test varsayımları yerine getiriliyor | Parametrik test varsayımları yerine getirilmiyorsa | |
| Bağımsız grup | Tek grup | Kitle ortalamasının önemlilik testi (One –Sample T test) | İşaret testi | Ki-kare (Chi-Square) |
| | 2 grup | İki ortalama arasındaki farkın önemlilik testi (Independent Sample T test) | Mann-Whitney U testi | •Beklenen değer 5'in üstünde ise 2x2 ki-kare •Beklenen değer 5'in altında ise Fisher's Exact test |
| | 3 veya daha fazla grup | Tek yönlü varyans (One-way ANOVA) | Kruskal-wallis testi | Çok gözlü Ki-kare |
| Bağımlı grup | 2 grup | İki eş arasındaki farkın önemlilik testi (Paired- Sample T test) | Wilcoxon testi | Mc Nemar testi |
| | 3 veya daha fazla grup | Tekrarlı ölçümlerde varyans analizi Repeated-measures ANOVA | Friedman test | Cochran Q testi |

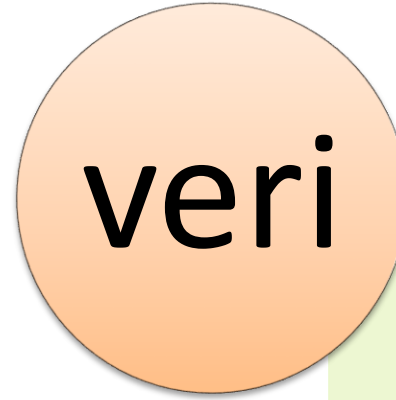
Verinin özelliđi

Arařtırmacı elde ettiđi verilerin özelliđini tanımlayabilmelidir.



parametrik

Hemoglobin,
homotokrit,
boy, kilo vs.



Non-parametrik

Cinsiyet,
meslek,
oturulan bölge
vs

Örnek Büyüklüğü (gruplardaki denek sayısı)

Araştırmaya alınacak bireylerin sayısı arttıkça istatistiksel testin gücü ve güvenilirliği artar.

İki ya da daha fazla grupta çalışılacaksa gruptaki birey sayılarının eşit seçilmesi standart yanlılığı azaltır.

Ölçümsel (parametrik testlerin) testlerin uygulanabilmesi için her grupta ki denek sayısının 30'un üzerinde olması gerekir.

Gruplardaki denek sayısı 30'un altında ise non-parametrik testlerden uygun olanı kullanılır.

İncelenen grupların bağımlı ya da bağımsız olması

- **Bağımlı grup (Bağımlı örneklem):** Aynı örnek birimlerden (aynı gruptaki kişilerden) bir değişkene ilişkin en az iki veri dizisi elde edildiğinde bu grup yapısına bağımlı grup (bağımlı örneklem) adı verilir. Aynı birimlerden iki ölçüm elde edilmişse bağımlı iki örneklem(bağımlı iki grup), üç ve daha fazla ise Bağımlı k-örneklem adı verilir.
- **Örnek:** A ilacının hipertansiyon üzerindeki etkisini inceleyen bir çalışmada 100 kalp hastasına hipertansiyon ilacı verilmeden önce ve verilmeye başlandıktan 10 gün sonra kan basıncı değerleri ölçülüyor.
 - Bu örnekte aynı grup üzerinde iki ölçüm yapılıyor. Dolayısı ile karşılaştırma bağımlı grup içinde yapılıyor.

İncelenen grupların bağımlı ya da bağımsız olması

- **Bağımsız grup (Bağımsız örneklem):** Bir değişken ile ilgili en az iki veri dizisi en az iki ayrı örnek birimlerden elde edildiğinde bu veri dizisine bağımsız grup (bağımsız örneklem) adı verilir. . Farklı iki gruptan iki ölçüm elde edilmişse bağımsız iki örneklem(bağımsız iki grup), üç ve daha fazla ise Bağımsız k-örneklem adı verilir.
- **Örnek:** A ilacının hipertansiyon üzerindeki etkisinin cinsiyete göre değişip değişmediğini inceleyen bir çalışmada 50 kadın ve 50 erkek kalp hastasına hipertansiyon ilacı verilmeden önce ve verilmeye başlandıktan 10 gün sonra kan basıncı değerleri ölçülüyor.
 - Bu örnekte ayrı iki grup üzerinde iki ölçüm yapılıyor. Dolayısı ile karşılaştırma bağımsız iki grup arasında yapılıyor.

Bağımlı grup

GRUP I → Ölçüm I → Ölçüm II

Ölçüm I ile Ölçüm II arasında fark var mı?

Bağımsız grup

GRUP I → Ölçüm

GRUP II → Ölçüm

iki ölçüm arasında fark var mı?

PARAMETRİK TEST VARSAYIMLARI

Veri ölçümle elde edilmiş olmalı

Gruplarda ki kişi (denek) sayısı 30'dan fazla olmalı


Veriler normal dağılıma uymalı

Gruplardaki kişi sayısı eşit ya da birbirine yakın olmalıdır.

Varyanslar eşit olmalı

Önemlilik testi aşamaları


1. Verinin özelliği nedir? (parametrik?, nonparametrik?)




2. Gruplar bağımlı mı? Bağımsız mı?



3. Gruplarda kaç gönüllü var?



4. Veriler normal dağılıyor mu?



5. İncelenen grupların bağımlı ya da bağımsız olması

| Gruplar | | Veri Tipi | | |
|---------------|------------------------|---|---|---|
| | | Ölçümle elde edilmiş veri (Parametrik veri) | | Sayımla elde edilmiş veri (Nonparametrik veri) |
| | | Parametrik test varsayımları yerine getiriliyor | Parametrik test varsayımları yerine getirilmiyorsa | |
| Bağımsız grup | Tek grup | Kitle ortalamasının önemlilik testi (SPSS >> Analyze>>Compare Means>> One – Sample T test) | İşaret testi | Ki-kare (SPSS>>Analze>>Nonparametric test>> Chi-Square) |
| | 2 grup | İki ortalama arasındaki farkın önemlilik test (SPSS>>Analyze>>Compare Means>> Independent Sample T test) | Mann-Whitney U testi (SPSS>>Analyze>>Nonparametric test>> 2 Independent Sample test- Mann-Whitney U testi) | •Beklenen değer 5'in üstünde ise 2x2 ki-kare •Beklenen değer 5'in altında ise Fisher's Exact test (SPSS>>Analyze>>Descriptive Statistic>>Crostabs>>Statistic>>Chi-Square) |
| | 3 veya daha fazla grup | Tek yönlü varyans (SPSS>>Analyze>>Compare Means>>One-way ANOVA) | Kruskal-wallis testi (SPSS>>Analyze>>Nonparametric test>> K Independent samples>>Kruskal-wallis testi) | Çok gözlü Ki-kare (SPSS>>Analyze>>Descriptive Statistic>>Crostabs>>Statistic>>Chi-Square) |
| Bağımlı grup | 2 grup | İki eş arasındaki farkın önemlilik testi SPSS>>Analyze>>Compare Means>> Paired-Sample T test) | Wilcoxon testi (SPSS>>Analyze>>Nonparametric test>>Two related samples>>Wilcoxon) | Mc Nemar testi (SPSS>>Analyze>>Nonparametric test>>Two related samples>>Mc Nemar) |
| | 3 veya daha fazla grup | Tekrarlı ölçümlerde varyans analizi Repeated-measures ANOVA (SPSS >> Analyze>> General Linear Model >> Repeated Measures) | Friedman test (SPSS>>Analyze>>Nonparametric test>>K related samples>>Friedman) | Cochran Q testi (SPSS>>Analyze>>Nonparametric test>>K related samples>> Cochran Q) |

Problem

- Lise öğrencilerinin ilk yardım bilgisini geliştirmede Akran liderli ilkyardım eğitimi yetişkin liderli ilk yardım eğitiminden daha etkili midir?

Amaç

- Lise öğrencilerine verilen akran ve yetişkin liderli ilk yardım eğitimlerinin etkisini değerlendirmek.

Hipotez

- Akran grubundaki öğrencilerin akran eğitimi sonrası ilkyardım bilgi puanları eğitim öncesinden daha yüksektir. (H1 tek yönlü).
- Öğrencilerin yetişkin eğitimi öncesi ve sonrası ilk yardım bilgi puanları arasında fark yoktur (H_0).

Tablo: Yetişkin ve Akran Liderli grupların eğitim öncesi ve sonrası ilk yardım bilgi puanlarının karşılaştırılması

| Gruplar | Eğitim öncesi | | Eğitim Sonrası | | t | p |
|------------------------------|---------------|------|----------------|------|------|-------|
| | X±SS | Sh | X±SS | Sh | | |
| Yetişkin Liderli Grup (n=60) | 60±5,38 | 0,76 | 61±4,72 | 0,66 | 2.14 | 0.14 |
| Akran Liderli Grup (n=60) | 60 ± 4,26 | 0,60 | 75± 5,58 | 0,78 | 3.14 | 0.000 |
| t | 1.13 | | 3.12 | | | |
| p | 0.670 | | 0.000 | | | |

Karar

- Akran grubundaki öğrencilerin akran eğitimi sonrası ilkyardım bilgi puanları eğitim öncesinde ki bilgi puanından 0.001 yanılma düzeyinde istatistiksel olarak daha yüksek bulunmuştur ($p < .001$).
- Öğrencilerin yetişkin Liderli eğitim öncesi ve sonrası ilk yardım bilgi puanları arasında 0.05 yanılma düzeyinde fark bulunmamıştır ($p > .05$)
- Yani Lise öğrencilerinin ilk yardım bilgisini geliştirmede akran liderli ilkyardım eğitimi yetişkin liderli ilk yardım eğitiminden daha etkilidir.

Problem

- **Gebenin çalışma durumu gebeliğe uyumu etkiler mi?**

Amaç

- **Çalışan ve çalışmayan geberin gebeliğe uyumlarını değerlendirmek.**

Hipotez

- **H1. Çalışan ve çalışmayan gebelerin gebeliğe uyum puanları arasında fark vardır (H1 çift yönlü)**
- **H2.Çalışan gebelerin gebeliğe uyum düzeyleri çalışmayan gebelerden daha yüksektir (H1 tek yönlü)**
- **H3. Çalışan ve çalışmayan gebelerin gebeliğe uyum puanları arasında fark yoktur (H_0).**

Gebelerin çalışma durumuna göre gebeliğe uyumlarının karşılaştırılması (n=480)

| Çalışma durumu | n | $X \pm Ss$ | test | p |
|----------------|-----|-------------------|------|-------|
| Çalışan | 117 | 142,68 \pm 5.48 | 3,64 | 0.014 |
| Çalışmayan | 273 | 150,09 \pm 6.23 | | |

H1. Çalışan ve çalışmayan gebelerin gebeliğe uyum puanları arasında fark vardır (H1 çift yönlü)

- **Karar:** Çalışan ve çalışmayan gebelerin gebeliğe uyum puanları arasında istatistiksel olarak fark vardır ($p < 0.05$)

H2. Çalışan gebelerin gebeliğe uyum düzeyleri çalışmayan gebelerden daha yüksektir (H1 tek yönlü)

- **Karar:** Çalışan gebelerin uyum düzeyleri çalışmayan gebelerden istatistiksel olarak daha yüksek bulunmuştur ($p < 0.05$)

Problem

- Kadınlarda ruhsal problem sıklığı erkeklerden daha fazla mıdır?

Amaç

- Kadın ve erkeklerde ruhsal problem sıklığını belirlemek.

Hipotez

- H1. Kadın ve erkeklerde ruhsal problem sıklığı farklıdır (H1 çift yönlü)
- H2. Kadınlarda ruhsal problem sıklığı erkeklerden daha fazladır (H1 tek yönlü)
- H3. Kadın ve erkeklerde ruhsal problem sıklığı farklı değildir (H_0).

Araştırmacı 972 erişkin bireyden GSA-28 ölçeğini kullanarak veri toplamıştır. GSA-28 ruhsal problemi belirlemeye yarayan bir tarama aracıdır. Ölçekten 5 ve üzerinde puan alanlarda ruhsal problem var denilmektedir. Araştırmaya 802 kadın, 170 erkek katılmıştır. Kadınların 271'inde, erkelerin 37'sinde ruhsal problem tespit edilmiştir.

Cinsiyete göre ruhsal problem sıklığının karşılaştırılması

| | Ruhsal problem yok | Ruhsal problem var | Toplam | Test | p |
|--------|--------------------|--------------------|--------|------|-------|
| Kadın | 531 (547,9) (a) | 271 (254,1) (b) | 802 | 6.73 | 0.002 |
| Erkek | 133 (116,1) (c) | 37 (53,9) (d) | 170 | | |
| Toplam | 664 | 308 | 972 | | |

Beklenen değer (b)

972 kişiden 308 kişide ruhsal problem var ise
802 kişiden kaçında ruhsal problem beklenir.

$$\text{Beklenen değer (b)} = \frac{802 * 308}{972} = 254,1$$

H1. Kadın ve erkeklerde ruhsal problem sıklığı farklıdır(H1 çift yönlü)

- **Karar:** Kadın ve erkeklerde ruhsal problem sıklığı istatistiksel farklı bulunmuştur($p < 0.05$)

H2. Kadınlarda ruhsal problem sıklığı erkeklerden daha fazladır(H1 tek yönlü)

- **Karar:** Kadınlardaki ruhsal problem sıklığı erkeklerden istatistiksel olarak daha yüksek bulunmuştur ($p < 0.05$)

Problem

- Kanser hastalarının yakınlarına verilen eğitim beden kitle indeksinin azaltılması üzerine etkili midir?

Amaç

- Kanser hastalarının yakınlarına verilen eğitim beden kitle indeksinin azaltılması üzerindeki etkisini değerlendirmek.

Hipotez

- H1. Deney grubundaki hasta yakınlarının eğitim sonrasındaki BKİ değerleri eğitim öncesinden daha düşüktür (H1 tek yönlü)
- H2. Kontrol grubundaki hasta yakınlarının eğitim öncesi ve sonrası BKİ değerleri arasında fark yoktur(H_0).

Deney ve kontrol grubunun ön test ve son test BKİ değerlerinin karşılaştırması

| Gruplar | Eğitim öncesi | | Eğitim Sonrası | | Test | p |
|-------------------------|---------------|------|----------------|------|-------|-------|
| | X±SS | Sh | X±SS | Sh | | |
| Deney grubu (n=50) | 26,17±5,38 | 0,76 | 26,00±4,72 | 0,66 | 0,690 | 0,494 |
| Kontrol grubu (n=50) | 26,11 ± 4,26 | 0,60 | 25,73± 5,58 | 0,78 | 1,074 | 0,288 |

H1. Deney grubundaki hasta yakınlarının eğitim sonrasındaki BKI değerleri eğitim öncesinden daha düşüktür (H1 tek yönlü)

- **Karar:** Deney grubundaki bireylerin eğitim öncesi ve sonrası ortalama BKI değerleri arasında istatistiksel olarak fark bulunmamıştır ($p > 0,05$).

H2. Kontrol grubundaki hasta yakınlarının eğitim öncesi ve sonrası BKI değerleri arasında fark yoktur (H_0).

- **Karar:** Kontrol grubundaki bireylerin eğitim öncesi ve sonrası ortalama BKI değerleri arasında istatistiksel olarak fark bulunmamıştır ($p > 0,05$).

SONUÇ

Kanser hastalarının
yakınlarına verilen eğitim
beden kitle indeksinin
azaltılması üzerine etkili
olmamıştır.

Problem

- Bir ortaokulda ağrı ile okul hemşiresine başvuran kız ve erkek öğrencilerin ağrı şiddeti farklı mıdır?

Amaç

- Bir ortaokulda ağrı ile okul hemşiresine başvuran kız ve erkek öğrencilerin ağrı şiddeti değerlendirmek.

Hipotez

- H1. Kız ve erkek öğrencilerin ağrı şiddeti arasında fark vardır (H1 çift yönlü)
- H2. Kız öğrencilerin ağrı şiddeti erkek öğrencilerden daha yüksektir (H1 tek yönlü)
- H3. Kız ve erkek öğrencilerin ağrı şiddeti arasında fark yoktur (H_0).

Cinsiyete göre ağrı şiddetinin karşılaştırılması

| Cinsiyet | n | $X \pm Ss$ | test | p |
|----------|-----|-----------------|------|------|
| Kız | 300 | 5.55 ± 2.15 | 1.76 | 0.07 |
| Erkek | 301 | 5.93 ± 2.30 | | |

~~H1. Kız ve erkek öğrencilerin ağrı şiddeti arasında fark vardır (H1 çift yönlü)~~

~~H2. Kız öğrencilerin ağrı şiddeti erkek öğrencilerden daha yüksektir (H1 tek yönlü)~~

H3. Kız ve erkek öğrencilerin ağrı şiddeti arasında fark yoktur (H_0).

- **Karar:** Kız ve erkek öğrencilerin ağrı şiddeti arasında istatistiksel olarak fark bulunmamıştır ($p > 0,05$).

Problem

- Kanser hastalarının yakınlarına verilen eğitim KKMM yapma davranış üzerine etkilimidir.?

Amaç

- Kanser hastalarının yakınlarına verilen eğitim KKMM yapma davranış üzerine etkisini değerlendirmek.

Hipotez

- H1. Deney grubunda eğitim sonunda KKMM yapan kişi sayısı eğitim öncesinden daha yüksektir (H1 tek yönlü)
- H2. Kontrol grubunda KKMM yapan kişi sayısı eğitim öncesi ve sonrasında farklı değildir (H_0).

Tablo 6.16. Deney ve kontrol gruplarının her ay kendi kendine meme muayenesi yapma durumlarının karşılaştırması

| | ÖN TEST | SON TEST | | | | | |
|----------------------------------|------------|----------|-------|-------------|---------|-------|----------------|
| | | DENEY | | | KONTROL | | |
| | | Evet | Hayır | χ^2 /p | Evet | Hayır | χ^{2*} /p |
| KKMM YAPMA DURUMU | Evet | 8 | 0 | ,741/,000 | 16 | 0 | 25,09/1,00 |
| | Hayır | 14 | 4 | | 0 | 13 | |

* Mc Nemar testi

H1. Deney grubunda eğitim sonunda KKMM yapan kişi sayısı eğitim öncesinden daha yüksektir (H1 tek yönlü)

- **Karar:** Deney grubunda eğitim sonunda KKMM yapan kişi sayısı eğitim öncesinden istatistiksel olarak daha yüksektir ($p < .001$).

H2. Kontrol grubunda KKMM yapan kişi sayısı eğitim öncesi ve sonrasında farklı değildir (H_0).

- **Karar:** Kontrol grubunda KKMM yapan kişi sayısı eğitim öncesi ve sonrasında istatistiksel olarak farklı değildir ($p > 0,05$).

SONUÇ

Kanser hastalarının
yakınlarına verilen eğitim
KKMM yapan kişi sayısının
arttırılmasında etkili
olmuştur.

Problem

- Sebze-meyve dostu program günde 5 porsiyon sebze meyve tüketenlerin sayısının arttırılmasında etkilimidir?

Amaç

- Sebze-meyve dostu programın etkisini değerlendirmek.

Hipotez

- H1. Son testte günde 5 porsiyon SM tüketenlerin sayısı ön testten daha yüksektir(H1 tek yönlü)
- H2. Son test ve izlem testte günde 5 porsiyon SM tüketenlerin sayısı farklı değildir (H_0).

Öğrencilerin ön test, son test ve izlem testlerdeki günde en az 5 porsiyon sebze ya da meyve tüketme

durumlarına göre karşılaştırılması

| | Ön-test | | Son-test | | İzlem test | | Q*/p | (a-b) χ^{2**}/p | (a-c) χ^{2**}/p | (b-c) χ^{2**}/p |
|--|---------|------|----------|------|------------|------|-----------------|-------------------------|-------------------------|-------------------------|
| | (a) | | (b) | | (c) | | | | | |
| | n | % | n | % | n | % | | | | |
| Günde 5 porsiyondan az sebze ve meyve tüketenler | 297 | 42,3 | 211 | 30,1 | 192 | 27,4 | 59,14 / .001 | 32,84 / .000 | 44,51/.00 0 | 1,77/.183 |
| Günde 5 porsiyondan fazla sebze ve meyve tüketenler | 405 | 57,7 | 491 | 69,9 | 510 | 72,6 | | | | |

*Cochran Q test

** Mc nemar test

H1. Son testte günde 5 porsiyon SM tüketenlerin sayısı ön testten daha yüksektir(H1 tek yönlü)

- **Karar:** Son testte günde 5 porsiyon SM tüketenlerin sayısı ön testten istatistiksel olarak daha yüksektir[$p < .001$].

H2. Son test ve izlem testte günde 5 porsiyon SM tüketenlerin sayısı farklı değildir (H_0).

- **Karar:** Program bitiminde ve program bittikten 3 ay sonraki izlemde günde en az 5 porsiyon meyve/sebze tüketme durumları arasında istatistiksel olarak fark bulunmamıştır($p > .05$).

SONUÇ

Program günde en az 5 porsiyon meyve/sebze tüketenlerin sayısının arttırılması üzerine etkili olmuştur.