İki Nicel Değişkenli Verileri İnceleme, Yanlılıklar ve İstatistiksel Yorumlama

MAT.11.1.2. Başkaları tarafından oluşturulan iki nicel değişkenli verilerin ilişkililiğine dayalı istatistiksel sonuç veya yorumları tartışabilme

HATIRLAYALIM

- Korelasyon-nedensellik yanılgısı; iki değişken arasında matematiksel bir bağ bulunmasının doğrudan bir sebep-sonuç ilişkisi anlamına gelmediğini unutan mantık hatasıdır.
- Örneklem yanlılığı; verilerin evreni temsil etmeyen taraflı bir gruptan toplanması sonucu analizlerin yanıltıcı çıkmasıdır.
- Ekstrapolasyon hatası; doğrusal modelin eldeki veri sınırlarının çok ötesindeki aşırı uç değerler için tahmin yapmada körü körüne kullanılmasıdır.

1 EŞLEŞTİRME

İstatistikte sık düşülen yanılgı ve hata türlerini tanımlarıyla eşleştiriniz.



- ☐ 1. Korelasyon Yanılgısı
- ☐ 2. Örneklem Yanlılığı
- ☐ 3. Ekstrapolasyon Hatası
- ☐ 4. Aykırı Değer Etkisi

a) Kümeden uzak tek bir noktanın korelasyon katsayısını yapay saptırması

b) İlişki yüksek çıktı diye birinin diğerine sebep olduğunu varsaymak

c) Veri toplanan grubun ana kütleyi temsil etmeyen taraflı yapıda olması

d) Gözlem aralığının çok dışındaki aşırı uçlar için hatalı tahmin yürütme

2 ÇOKTAN SEÇMELİ

Görseldeki 0-10 yaş arası çocukların boy ve yaş verisine dayalı regresyon doğrusu (y = 6x + 75) kullanılarak bir kişinin 50 yaşındaki boyunun 375 cm olarak tahmin edilmesi hangi hatadır?



(A) Negatif korelasyonu pozitif yönlü okuma hatası

(B) Örneklem kümesindeki medyan değerini yanlış hesaplama hatası

(C) Modeli veri aralığının çok ötesine uygulayan ekstrapolasyon hatası

(D) Standart sapma değerini sıfırdan küçük kabul etme hatası

3 ÇOKTAN SEÇMELİ

Bir gazete haberinde 'Kahve tüketimi ile gelir düzeyi arasında güçlü pozitif ilişki bulundu; bu yüzden günde üç fincan kahve içmek insanı zengin yapar' denilmiştir. Bu haberdeki temel istatistiksel mantık hatası nedir?



(A) Korelasyonu nedensellik gibi yorumlayıp üçüncü değişkenleri görmezden gelme hatası

(B) Grafik çiziminde dikey eksenin yatay eksenden daha uzun gösterilmesi hatası

(C) Kahve fiyatlarının borsa döviz kurlarından bağımsız belirlenmesi hatası

(D) Verilerin sadece tek bir kahve türü üzerinden hesaplanması hatası

4 BOŞLUK DOLDURMA

Boş bırakılan yere uygun kavramı yazınız.



nedensellik

1. İki nicel değişken arasında güçlü matematiksel ilişki bulunsa dahi biri diğerinin kesin sebebi olmak zorunda değildir; bu duruma _____ yanılgısı denir.

5 BOŞLUK DOLDURMA

Boş bırakılan yere uygun kavramı yazınız.

interpolasyon

1. Eldeki veri kümesinin sınırları içerisinde kalarak ara değerler için yapılan tahmine _____ (içdeğerbiçim) denir.

6 ÇOKTAN SEÇMELİ

Bir anket şirketinin 'Halkın teknolojiye bakışı' araştırmasını sadece bir yazılım şirketinin çalışanlarına uygulayarak genel bir rapor yayımlaması hangi istatistiksel kusurdur?

(A) Eksenlerin logaritmik ölçekte çizilmemiş olması

(B) Toplumun bütününü temsil etmeyen yanlı bir örneklem seçilmiş olması

(C) Katılımcıların tamamının aynı şehirden seçilmiş olması

(D) Anket sorularının internet üzerinden yanıtlanmış olması

7 EŞLEŞTİRME

İstatistiksel analiz adımlarını taşıdıkları doğrulama işlevleriyle eşleştiriniz.

VERİ VE MODEL DENETİMİ KONTROL LİSTESİ TABLOSU

Kontrol Ölçütü	Sorulacak Kritik Soru	Doğrulama Amacı
Örneklem Temsili	Veri kimlerden toplandı?	Yanlılığı Tespit Etme
Korelasyon Bağı	Üçüncü bir etken var mı?	Nedensellik Tuzağı
Model Sınırı	Tahmin aralık içinde mi?	Ekstrapolasyonu Önleme

☐ 1. Serpilme Grafiği

☐ 2. Korelasyon Katsayısı (r)

☐ 3. Artık (Kalıntı) Analizi

☐ 4. Alan Kısıtı Belirleme

a) Doğrusal modelin geçerli olduğu güvenilir gözlem aralığını çizme

b) İlişkinin sayısal yönünü (-1 ile +1 arası) ve derecesini ölçme

c) Veri noktalarının doğrusal bir eğilim gösterip göstermediğini görme

d) Modelden sapan gerçek değerlerin dağılımını kontrol etme

8 ÇOKTAN SEÇMELİ

Serpilme diyagramında yer alan sekiz veri noktasının yedisi r = +0,90 eğiliminde iken tek bir uç noktanın katsayıyı r = +0,35'e düşürmesi karşısında ne yapılmalıdır?

AYKIRI (UÇ) DEĞERİN MODELİ SAPTIRMA ETKİSİ

Aykırı Değersiz Model	Tek Aykırı Değer Eklenince
r = +0,90 Güçlü İlişki	r = +0,35 Yapay Düşüş
Gerçekçi Doğrusal Eğilim	Modeli Ciddi Yanıltır

(A) Bütün veriler çöpe atılarak araştırmadan tamamen vazgeçilmelidir

(B) Grafiğin eğim açısı yapay olarak kırk beş dereceye yükseltilmelidir

(C) Kalan yedi noktanın tamamı da grafikten derhal silinmelidir

(D) O uç noktanın ölçüm hatası veya özel bir durum olup olmadığı incelenmelidir

9 ÇOKTAN SEÇMELİ

Bir kentte sokak lambalarının sayısı ile kütüphanelerdeki kitap sayısı arasında r = +0,88 pozitif ilişki görülmüştür. Bu iki değişken arasındaki ortak gizli etken nedir?

ELEŞTİREL VERİ OKURYAZARLIĞI BASAMAKLARI

1. Kaynak Doğrulama	2. Eksen Ölçeği	3. Değişken Bağı	4. Sonuç Analizi
Veriyi kim üretti?	Sıfırdan mı başlıyor?	Mantıklı bir ilişki mi?	Abartılı genelleme var mı?

(A) Şehrin toplam nüfusunun ve yerleşim alanının büyümesi değişkeni

(B) Sokak lambalarının kitap okumayı gece boyunca zorunlu kılması

(C) Kütüphanelerdeki kitapların elektrik enerjisiyle basılması

(D) Şehirdeki yağış miktarının her yıl düzenli olarak düşmesi

10 ÇOKTAN SEÇMELİ

Başkaları tarafından yayımlanan bir istatistiksel grafiği eleştirel gözle incelerken hangisine dikkat etmek en öncelikli bilimsel tutumdur?

(A) Verilerin güvenilir bir kaynaktan toplanıp toplanmadığına ve yanlılık olup olmadığına

(B) Grafiğin hangi renk tonlarıyla basıldığına ve gölgelendirme efektlerine

(C) Yazı puntosunun diğer ders kitaplarıyla aynı boyutta olup olmadığına

(D) Rapordaki matematiksel denklemlerin en az on satır uzunluğunda olmasına

11 BOŞLUK DOLDURMA

Boş bırakılan yere uygun kavramı yazınız.

İSTATİSTİKSEL AKIL YÜRÜTME DÖNGÜSÜ

Aşama 1	Aşama 2	Aşama 3	Aşama 4
Veri İnceleme	Yanlılık Arama	Sebep Sorgulama	Sağlam Sonuç

kalıntı

1. Model denkleminde gözlenen gerçek değer ile modelin tahmin ettiği değer arasındaki farka _____ (hata) denir.

12 AÇIK UÇLU

İstatistikte iki değişken arasında yüksek korelasyon çıkmasının neden her zaman nedensellik anlamına gelmediğini açıklayınız.

derskutusu.com · © Ders Kutusu · Tüm hakları saklıdır.