



FİNAL SINAV KAĞIDI

Adı:	Dersin Adı: DENEY TASARIMI	Not
Soyadı:	Dersin Kodu: IST3002	
Numarası:	Bölümü: İSTATİSTİK	
İmzası:	Son Yükleme Tarihi: 01/07/2020 Saat 15:00	

Açıklamalar

1. Cevap kağıdınızın her birine ad, soyad, okul numarası yazınız ve imza atınız.
2. Sisteme yüklediğiniz PDF dosyasının ismini "Ad Soyad Okul Numarası" olarak düzenleyiniz.
3. Sınav ile ilgili problemlerinizi için sınav süresince fatih.kizilaslan@marmara.edu.tr e-posta adresinden iletişime geçebilirsiniz.
4. Bu ödev kişisel başarınızı göstereceğinden ödevin cevaplarını bu ders ile ilgili kendi bilgilerinizi kullanarak yardım almadan yapmalısınız.
5. Tüm cevaplarınızı anlaşılır bir biçimde açıklayarak yazınız. Açıklaması olmayan cevaplar değerlendirilmeyecektir.

SORULAR

A) (75 puan) Bu bölümdeki sorular uygulama ile ilgilidir. Cevaplarınızı R Markdown kullanarak oluşturunuz. Sisteme yüklemeyen önce PDF formatına çeviriniz veya Word formatına çevirerek sonrasında PDF yapınız.

Ayrıca, sınav sonunda ile ilgili R Markdown kodunuzu isim_soyisim olarak kaydederek e-posta ile "fatih.kizilaslan@marmara.edu.tr" adresine gönderiniz.

Kaggle'da "https://www.kaggle.com/anakitbarai507/breast-cancer-dataset" adresinde yer alan (ayrıca BYS'de bulunan e-posta adreslerinize gönderdiğim "datasets_631294_1123441_dataR2.csv") göğüs kanseri ile ilgili veriyi kullanarak aşağıdaki soruları cevaplayınız. (Anlamlılık düzeyi $\alpha = 0.05$ olarak alınacaktır.)

1. Veride bulunan Age değişkenini 5 veya 6 düzeyli ve BMI değişkenini 3 veya 4 düzeyli olacak biçimde kendi istediğiniz gibi rastgele bir biçimde düzeylere ayırarak faktör olarak tanımlayınız.

(Not: Ayırdığımız düzeylerdeki toplam gözlem sayısının verideki toplam gözlem sayısına eşit olduğunu kontrol ediniz.)

Örneğin,

Age	Faktör Düzeyi	BMI	Faktör Düzeyi
$24 \leq \text{Age} < 32.5$	1	$18 \leq \text{BMI} < 21.5$	1
$32.5 \leq \text{Age} < 45$	2	$21.5 \leq \text{BMI} < 25$	2
$45 \leq \text{Age} < 57$	3	3
.....	4	4
.....	5		
.....	6		

Classification değişkeninin 1 (sağlıklı) ve 2 (hasta) olmak üzere iki düzeyi vardır. Bu değişkeni de faktör olarak tanımlayınız.

2. Bağımlı değişken olarak ayrı ayrı Glucose, Insulin, HOMA, Leptin ve Resistin değişkenlerini ve faktör olarak Classification değişkenini kullanarak varyans analizlerini yapınız. (Not: 5 farklı analiz yapılacaktır.)

a) Yaptığımız bir analiz için uygun modeli yazınız. (A4 kağıdına yazılabilir.)

b) Her bir analiz sonucunu 2 cümle ile açıklayarak yorumlayınız.

c) Resistin için yaptığımız analizde varsayımların kontrolünü yapınız ve sonuçlarını birer cümle ile yorumlayınız.

3. Yukarıda tanımlandığı gibi yeni oluşturduğunuz **Age ve BMI faktörlerini kullanarak**
- Bağımlı değişken: **Resistin**, Faktörler: **Classification, Age**
 - Bağımlı değişken: **Resistin**, Faktörler: **Classification, BMI** olmak üzere varyans analizlerini (etkileşimi de dikkate alarak) yapınız.
 - Bu analizler için uygun modelleri yazınız. (**A4 kağıdına yazılabilir.**)
 - Analiz sonuçlarını birkaç cümle ile yorumlayınız.
 - b' de yaptığımız analiz için varsayımların kontrolünü yapınız ve sonuçlarını birer cümle ile yorumlayınız.
4. 2. ve 3. sorularda yaptığınız analizlere göre en iyi model hangisidir ? Açıklayınız.

B) (25 puan) Bu bölümdeki soruların cevaplarını bir (1) sayfa A4 kağıdına yazarak çözünüz.

Değişim kaynağı	Kareler toplamı	Serbestlik derecesi	Kareler ortalaması	F-test değeri
Denemeler	1010.56	4	?	?
Bloklar	?	?	?	?
Hata	169.33	20	?	
Toplam	1503.71	29		

Faktör ve blokların düzeyleri **sabit etkili** olmak üzere yukarıda verilen varyans analiz tablosuna göre aşağıdaki soruları cevaplayınız.

- Bu analiz için uygun modelin adını ve modeli yazınız.
- Tablodaki boşlukları doldurunuz. $\alpha = 0.05$ anlamlılık düzeyinde uygun tüm hipotezleri yazınız ve test ediniz. Analizin sonuçlarını birkaç cümle ile yorumlayınız.
- Eğer faktörün düzeyleri rastgele seçilseydi ilgili hipotezi yazınız. Bu hipotezi de test ediniz ($\alpha = 0.05$ için) ve 2' de yaptığımız ile karşılaştırınız.

BAŞARILAR

Doç. Dr. Fatih KIZILASLAN