**Öğrencinin Adı-Soyadı:**

**Numarası:**

### Soru 1: Aşağıda i=1,2,...,n olmak üzere verilen

$$Y\_{i}=β\_{0}+β\_{1}X\_{i}+ε\_{i}$$

Basit Doğrusal Regresyon modeli hataların birbirleriyle ilişkisiz, sıfır ortalama ve bilinmeyen varyans  sahip olduklarını varsayar.

1. $E\left(Y\_{i}\right) $ ve $Var\left(Y\_{i}\right)$ bulunuz (10p)
2. $\sum\_{i=1}^{n}X\_{i}e\_{i}=0$ olduğunu gösteriniz. (10p)

**Soru 2:** İngiltere’de yapılan bir araştırmada 8 ay boyunca gelen turist sayısı ile ziyaretleri sırasında harcadıkları miktarlar (Sterlin) gözlenmiştir. Bu veriye göre;

|  |  |
| --- | --- |
| **Ziyaretçi sayısı(x1000)**  | **Harcama(x100000)£** |
| 2450 | 1370 |
| 2480 | 1350 |
| 2540 | 1400 |
| 2420 | 1330 |
| 2350 | 1270 |
| 2290 | 1210 |
| 2400 | 1330 |
| 2460 | 1350 |

1. X ile Y arasındaki korelasyon katsayısı $r\_{xy}=\frac{\sum\_{i=1}^{n}\left(Y\_{i}-\overbar{Y}\right)\left(X\_{i}-\overbar{X}\right)}{\sqrt{\sum\_{i=1}^{n}\left(Y\_{i}-\overbar{Y}\right)^{2}\sum\_{i=1}^{n}\left(X\_{i}-\overbar{X}\right)^{2}}} $ formülü ile hesaplandığına göre, bu iki değişken arasındaki korelasyonu hesaplayınız. (10p)
2. $Y\_{i}=β\_{0}+β\_{1}X\_{i}+ε\_{i} $ Modelinin veriye uygun olabileceğini nasıl açıklarsınız (5p)
3. Eğim ve kesişim parametrelerini tahmin ediniz ve bu tahmin değerlerini yorumlayınız (10p).
4. Tahmin denklemini yazınız. Varyans analizi tablosunu oluşturarak $α=0.05 $ için regresyon modelinin istatistiksel olarak anlamlı olup olmadığını test ediniz. (10p)
5. Bağımlı değişkenin kendi ortalaması etrafındaki değişiminin bağımsız değişken ile açıklanabilen kısmını belirleyiniz. (5p)
6. Eğim ve kesişim parametrelerinin %95’lik güven aralıklarını oluşturunuz. (15p)
7. Gelen turist sayısı 2600 olsaydı ortalama harcanan miktar ne olurdu tahmin ediniz. Bu ortalama değer için %95 lik güven aralığını oluşturunuz. (10p)
8. X açıklayıcı değişkeninin etkisinin anlamlılığını test ederek yorumlayınız. (5p)
9. $\hat{Y}\_{i}$ tahmin değerlerini hesaplayınız. (10p)

**Başarılar**

**Prof. Dr. Vedide Rezan USLU**

**Cevaplar**

**1.(a)** Betalar bilinmeyen parametreler ve X rastgele değişken olmadığı için aynı zamanda hatanın beklenen değeri sıfır varyansı $σ^{2}$ sabit olduğu varsayıldığından; Bağımlı değişkenin beklenen değeri

$$E\left(Y\_{i}\right)=E\left(β\_{0}+β\_{1}X\_{i}+ε\_{i}\right)=E\left(β\_{0}+β\_{1}X\_{i}\right)+E\left(ε\_{i}\right)$$

$$=β\_{0}+β\_{1}X\_{i}+E\left(ε\_{i}\right)=β\_{0}+β\_{1}X\_{i}$$

varyansı

$$Var\left(Y\_{i}\right)=Var\left(β\_{0}+β\_{1}X\_{i}+ε\_{i}\right)=Var\left(ε\_{i}\right)=σ^{2}$$

elde edilir.

**1.(b)**  $\sum\_{i=1}^{n}X\_{i}e\_{i}=0$ nin ispatı;

$$\sum\_{i=1}^{n}X\_{i}\left(Y\_{i}-\hat{Y}\_{i}\right)=0$$

$$\sum\_{i=1}^{n}X\_{i}\left(Y\_{i}-\left(\hat{β}\_{0}+\hat{β}\_{1}X\_{i}\right)\right)=0$$

$$\sum\_{i=1}^{n}X\_{i}Y\_{i}-\sum\_{}^{}X\_{i}\left(\hat{β}\_{0}+\hat{β}\_{1}X\_{i}\right)=0$$

$$\sum\_{i=1}^{n}X\_{i}Y\_{i}-\sum\_{}^{}X\_{i}\hat{β}\_{0}-\sum\_{}^{}X\_{i}\left(\hat{β}\_{1}X\_{i}\right)=0$$

$$\sum\_{i=1}^{n}X\_{i}Y\_{i}-\sum\_{}^{}X\_{i}\left(\overbar{Y}-\hat{β}\_{1}\overbar{X}\right)-\hat{β}\_{1}\sum\_{}^{}X\_{i}^{2}=0$$

$$\sum\_{i=1}^{n}X\_{i}Y\_{i}-\overbar{Y}\sum\_{}^{}X\_{i}+\hat{β}\_{1}\sum\_{}^{}X\_{i}\overbar{X}-\hat{β}\_{1}\sum\_{}^{}X\_{i}^{2}=0$$

$$\left\{\sum\_{}^{}X\_{i}Y\_{i}-\overbar{Y}\sum\_{}^{}X\_{i}\right\}-\hat{β}\_{1}\left\{\sum\_{}^{}X\_{i}^{2}-\overbar{X}\sum\_{}^{}X\_{i}\right\}=0$$

$$S\_{XY}-\hat{β}\_{1}S\_{XX}=0$$

$$S\_{XY}-\frac{S\_{XY}}{S\_{XX}}S\_{XX}=0$$

Tamamlanmış olur.

**2.(a)** Y=Turistlerin harcamaları, X=gelen turist sayısı olmalıdır. Buna göre

$r\_{xy}=0.962$ .

**2.(b)** Korelasyon katsayısı iki değişken arasındaki doğrusal ilişkinin şiddetini ve yönünü vermektedir. Bu örnekte de bu katsayı +1 e çok yakın bulunmuştur. Bu iki değişken arasında kuvvetli ve pozitif yönlü doğrusal bir ilişki vardır.

**2.(c)** En küçük kareler tahmin edicileri $\hat{β}\_{1}=\frac{S\_{XY}}{S\_{XX}}$ $\hat{β}\_{0}=\overbar{Y}-\hat{β}\_{1}\overbar{X}$ ile bulunmaktadır. Formülleri uyguladığımızda,

$\hat{β}\_{0}=-467$ $\hat{β}\_{1}=0,7399≈0.74$ bulunur. Turist sayısında bir birimlik (yani ziyaretçi sayısı 1000 kişi arttığında) artış, harcama miktarında 0,7399 poundluk (£739.9) bir artışa neden olur.

2.(d) Tahmin denklemi (regresyon denklemi) $\hat{Y}\_{i}=-467+0.74X\_{i}$ olarak elde edilmiştir. Varyans analizi tablosunu oluşturmak için kareler toplamlarının bulunması gereklidir. Kareler toplamları sırasıyla aşağıda hesaplanmıştır.

$$GKT=S\_{YY}=25187.5$$

$$RKT=\hat{β}\_{1}S\_{XY}=0.7399\*31512.5=23316.0987$$

$$AKT=GKT-RKT=25187.5-23316.0987=1871.4013$$

Varyans Analizi Tablosu

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Değişim Kaynağı | s.d. | KT | KO | FHata varyansının tahmin edicisi =AKO |
| Regresyon | 1 | 23316.0987 | 23316.0987 | 74.75 |
| Artık | 6 | 1871.4013 | 311.4012 $\left\{\hat{σ}^{2}\right\}$ |  |
| Genel | 7 | 25187.5 |  |  |

Hipotez

$$H\_{0}=Regresyon modeli anlamlı değildir$$

$$H\_{1}=Regresyon modeli anlamlıdır$$

Şeklinde kurulur. Bu hipotezi test etmek için hesaplanan F test istatistiğinin değeri 74.75 dir. Tablodan bulunacak olan kritik değer ise $F\_{1,6,0.05}=5.99$ dur. $F\_{hesap}=74.75>5.99=F\_{tablo}$ olduğundan sıfır hipotezi ret edilir. Yani Regresyon modeli bu veri için istatistiksel bakımdan anlamlı bir modeldir.

**2.(e)** Bu şıkta determinasyon (açıklama) katsayısı sorulmaktadır. $R^{2}=\frac{RKT}{GKT}=\frac{23316.0987}{25187.5}=0.925$ , $R^{2}×100=\%92.5$ olarak dönüştürülen bu katsayının yorumu: bağımlı değişken Y’nin içindeki toplam değişimin %92.5 luk kısmı bağımsız değişkendeki değişmelerle açıklanmaktadır.

**2. (f)** Regresyon katsayılarının standart hataları sırasıyla

$sh\left(\hat{β}\_{0}\right)=\sqrt{\hat{σ}^{2}\left(\frac{1}{n}+\frac{\overbar{X}^{2}}{S\_{XX}}\right)}$ $sh\left(\hat{β}\_{1}\right)=\sqrt{\frac{\hat{σ}^{2}}{S\_{XX}}}$

formülleri ile hesaplanır. $sh\left(\hat{β}\_{0}\right)=207.3497$ ve $sh\left(\hat{β}\_{1}\right)=0.0855 $bulunur. Aralık tahminleri

$β\_{0}$ için %95’lik aralık tahmini $\hat{β}\_{0}\mp sh\left(\hat{β\_{0}}\right)\*t\_{\left(n-2\right), 0.025}$

$β\_{1}$ için %95’lik aralık tahmini $\hat{β}\_{1}\mp sh\left(\hat{β\_{1}}\right)\*t\_{\left(n-2\right), 0.025}$

Tablodan $t\_{\left(8-2\right), 0.025}=2.447$ bulunur. Yerine yerleştirildiğinde

$$\left(-974.3847, 40.3847\right)$$

$$\left(0.53068, 0.949118\right)$$

Şeklinde aralık tahminleri bulunmuş olur.

**2. (g)** Gelen turist sayısı 2600 olsaydı ortalama harcanan miktar ne olurdu tahmin ediniz.

Yani $X\_{0}$=2600 olduğunda Y’nin ortalama tahmin değeri regresyon denkleminde yerine yerleştirilerek bulunur. $\hat{Y}\_{0}=-467+0.74\*2600=1457$ bulunur. Gelecek ziyaretçi sayısı 2600 olduğunda ortalama harcama miktarı 1457 olarak tahmin edilir.

Bu ortalama değer için %95 lik güven aralığı $\hat{Y}\_{0}\pm sh\left(\hat{Y}\_{0}\right)×t\_{n-2,0.025}$ ile bulunacaktır.

$$sh\left(\hat{Y}\_{0}\right)=\sqrt{\hat{σ}^{2}\left(\frac{1}{n}+\frac{\left(X\_{0}-\overbar{X}\right)^{2}}{S\_{XX}}\right)}$$

$$sh\left(\hat{Y}\_{0}\right)=\sqrt{311.4012\left(\frac{1}{8}+\frac{\left(2600-2423,75\right)^{2}}{42587,5}\right)}=16.3115$$

$$1457\pm 16.3115×2.447$$

$\left(1416.9915, 1497.0084\right)$ ziyaretçi sayısı 2600 olduğunda ortalama harcama %95 ihtimalle bu aralık içerisinde yer alır.

2. (h) Burada $H\_{0}:β\_{1}=0$ yokluk hipotezinin $H\_{1}:β\_{1}\ne 0$ alternatif hipotezine karşılık testi yapılmalıdır. Bu test için test istatistiği $t=\frac{\hat{β}\_{1}}{sh\left(\hat{β}\_{1}\right)}$ formülü ile bulunur. Buna göre test istatistiğinin değeri

$$t=\frac{\hat{β}\_{1}}{sh\left(\hat{β}\_{1}\right)}=\frac{0.7399}{0.0855}=8.6538$$

Bulunur.

$\left|t\right|=8.6538>2.447=t\_{\left(8-2\right), 0.025}$ olduğundan $H\_{0}$ hipotezi ret edilir. Yani X’in katsayısı olan $β\_{1}$ için bulunan tahmin değeri istatistik bakımdan önemlidir. Yani X değişkeni Y’nin değişimini açıklamada önemli bir değişkendir.

**2. (h)** Tüm X ler regresyon denkleminde yerine konur. Böylelikle tüm gözlenen değerlere karşılık gelen tahmin değerleri hesaplanır. Bunlar;

|  |
| --- |
| $$\hat{Y}\_{i}$$ |
| 1345,76 |
| 1367,95 |
| 1412,35 |
| 1323,56 |
| 1271,77 |
| 1227,37 |
| 1308,76 |
| 1353,15 |