

Regresyon analizi

Wikipedi, özgür ansiklopedi

Başlığın diğer anlamları için Regresyon sayfasına bakınız.

Regresyon analizi iki ya da daha çok değişken arasındaki ilişkiyi ölçmek için kullanılan analiz metodudur. Eğer tek bir değişken kullanılarak analiz yapılıyorsa buna tek değişkenli regresyon, birden çok değişken kullanılıyorsa çok değişkenli regresyon analizi olarak isimlendirilir. Regresyon analizi ile değişkenler arasındaki ilişkinin varlığı, eğer ilişki var ise bunun gücü hakkında bilgi edilinebilir. Regresyon terimi için öz Türkçe olarak**bağlanım** sözcüğü kullanılması teklif edilmiş ise de^[1] Türk ekonometriciler arasında bu kullanım yaygın değildir

Örneğin, bir *ziraat mühendisi* *buğday* verimi ve gübre miktarı arasındaki ilişkiyi, bir mühendis, *basınç* ve *sıcaklık*, bir ekonomist gelir düzeyi ve tüketim harcamaları, bir eğitimci öğrencilerin devamsızlık gösterdiği gün sayıları ve başan dereceleri arasındaki ilişkiyi bilmek isteyebilir. Regresyon, iki (ya da daha çok) değişken arasındaki doğrusal ilişkinin *fonksiyonel* şeklini, biri bağımlı diğeri bağımsız değişken olarak bir doğru denklemi olarak göstermekle kalmaz, değişkenlerden birinin değeri bilindiğinde diğeri hakkında kestirim yapılmasını sağlar. Genellikle bu iki (veya çok) değişkenlerin hepsinin niceliksel ölçekli olması zorunluluğu vardır

Regresyonda, değişkenlerden biri bağımlı diğerleri bağımsız değişken olmalıdır. Buradaki mantık eşitliğin solunda yer alan değişkenin sağında yer alan değişkenlerden etkilenmesidir. Sağda yer alan değişkenlerse diğer değişkenlerden etkilenmemektedir. Burada etkilenmemek matematiksel anlamda bu değişkenleri bir doğrusal denkleme koyduğumuzda etki yapması anlamındadır. *Çoklu doğrusallık* *ardışık bağımlılık* sorunları kastedilmemektedir

İçindekiler

Regresyon analizinin tarihi

Temeldeki varsayımlar

Doğrusal regresyon

Anakütle doğrusal regresyon modeli

İki Değişkenli regresyon katsayı kestirimleri

Çok değişkenli regresyon katsayı kestirimleri

Hatalar varyansı ve toplam kareler

Kestirim denklemin genel uyum iyiliğinin çıkarımsal kontrolü

Belirleme katsayısı yani R-kare (**R**²) değeri

F-testi

Kestirimi yapılan her tek regresyon parametresinin değerinin çıkarımsal kontrolü

İnterpolasyon ve ekstrapolasyon

Diğer yaklaşımlar

Ağırlıklı en küçük kareler yöntemi

Değişkenlerde-hatalar modeli

Genelleştirilmiş doğrusal model

Güçlü regresyon

Ayrık bağımlı değişken

Doğrusal olmayan regresyon

İşsel kaynaklar

Kaynakça

Bibliyografya

Dış bağlantılar

Regresyon analizinin tarihi

Regresyon yönteminin ilk şekli *en küçük kareler* prensibidir ve ilk olarak *Adrien Marie Legendre* tarafından 1805 yılında ortaya atılmıştır^[2]. Hemen takiben 1809 yılında *C.F. Gauss* ^[3] aynı yöntemi açıklamıştır. *En küçük kareler* terimi Legendre tarafından *moindres carrés* olarak kullanılmış, ancak Gauss aynı yöntemi 1795 yılından beri kullandığını iddia etmiştir. Legendre ve Gauss bu yöntemi astronomik gözlemlerden uydularının güneş etrafındaki yörüngelerini tespit etmek için kullanırken ortaya çıkartmışlardır. 1748 yılında *Eüler*'in aynı problem üzerinde uğraştığı, fakat başarı sağlayamadığı bilinmektedir. En küçük kareler kuramında sonraki gelişme Gauss'in 1821 yılında yayınladığı bir makalede ortaya çıkartılmış^[4] ve bu yayında Gauss sonradan kendi adı verilen *Gauss-Markov* teoreminin bir şeklini açıklamıştır.

Regresyon terimi 19. yüzyılda İngiliz istatistikçisi Francis Galton tarafından bir biyolojik inceleme için ortaya atılmıştır. Bu incelemenin ana konusu kalıtım olup, aile içinde baba ve annenin boyu ile çocukların boyu arasındaki bağlantıyı araştırmakta ve çocukların boylarının bir nesil içinde eski ata nesillerinin ortalamasına geri döndüklerini yani bir nesil içinde *ortalamaya geri dönüş* olduğu inceleme konusudur. Galton *geri dönüş* terimi için ilk yazısında İngilizce olarak *reversion* terimi kullanmışsa da sonradan aynı anlamda olan *regression* sözcüğü kullanmıştır^{[5],[6]}. Bu çalışmalarında Galton istatistiksel 'regresyon' kavramını ve yöntemini de geliştirmiştir. *Udny Yüle* ve *Karl Pearson* bu yöntemi daha geniş genel istatistiksel alanlara uygulayıp geliştirmişlerdir.^{[7][8]}. Bu yazılarda bağımlı ve bağımsız değişkenlerin normal dağılım gösterdiği varsayılmaktadır. Bu kısıtlayıcı varsayım *R.A. Fisher* 1922 ve 1925 yıllarındaki yayınları ile sadece bağımlı değişkenin koşullu dağılımının normal olduğu hallere uygulanmak üzere daha genişletilmiştir^{[9][10]}.

Bu kavramları ve yöntemleri genel olarak, kalıtım konusu dışında "ortalamaya geri dönüş" ile hiçbir ilgisi olmayan konularda, kullanan istatistikçiler *regresyon* terimini kullanmakta devam etmişlerdir. Zamanımızda, bu terim, kavram ve yöntemin Galton'un konusu ile bütün ilişkisi yok olmuştur ve artık*regresyon* terimi doğrusal bağlantı bulunması *veeğri uydurma* ile eş anlamlar vermektedir.

Temeldeki varsayımlar

Doğrusal regresyon yöntemini kullanmak için temelde şu varsayımların bulunduğu kabul edilmektedir^[1]

- Çıkarımsal yöntem olduğu için kullanılan iki değişkenli örneklemın ya istatistiksel rastgele örneklem olduğu ya da anakütleyi çok iyi temsil ettiği bilinmektedir
- Bağımlı değişken içinde hata bulunmaktadırBu hatanın bir *rassal değişken* olduğu ve *ortalama* hatanın sıfır olduğudur. *Sistematik hata* da bulunması mümkündür ama bu hatanın incelemeye alınması regresyon analizi kapsamı dışındadır

- Bağımsız değişken hatasızdır Eğer bağımsız değişkende hata bulunduğu varsayılırsa özel bir yöntem şekli olan değişkenler-içinde-hata model teknikler kullanılarak model kurulmalıdır.
- Hatalar zaman içinde ve kendi aralarında birbirine bağımlı değildir Bu otokorelasyon veya serisel korelasyon bulunmaması varsayımı adı verilir
- Hata varyansı sabittir ve veriler arasında hiç değişmediği varsayılır Bu eşvaryanslılık veya homoskedastisite varsayımı adı ile anılır Eger bu varsayım uygun değilse ağırlıklı en küçük kareler yöntemi uygulanabilir
- Hataların varyans-kovaryans matrisinin çapraz elamanları sabit hata varyansı olur ve matrisin diğer çapraz dışı elemanları 0 olur
- Eğer çoklu regresyon analizi yapılıyor ve uc veya daha çok parametre için kestirim isteniyorsa, bağımsız değişkenlerin birbirleri ile bağlantısının olmaması gerekir Bu çoklu doğrusallık (multicollinearity) olmaması varsayımı adı verilir
- Hatalar bir normal dağılım gösterir. Eğer bu hataların normalliği varsayımı uygun değilse genelleştirilmiş doğrusal model uygulanabilir.

Doğrusal regresyon

Anakütle doğrusal regresyon modeli

Doğrusal regresyonda, anakütle model belirlenmesine göre ^[1] bağımlı değişken ***y_i*** *parametrelerin bir doğrusal birleşigi* olur. Dikkat edilirse *parametrelerden* bahis edilmektedir, çünkü bağımsız değişkenlerin bir doğrusal bileşigi olması gerekli değildirÖrneğin, tek bir bağımsız değişkenli (***x_i***) ve iki parametrelili (***β₀*** ve ***β₁***):

$$\text{dogru: } y_i = \beta_0 + \beta_1 x_i + \epsilon_i, i = 1, n$$

Burada ***ε_i*** bir *hata terimi*dir ve ***i*** alt-indeksleri bir belirlenmiş mümkün gözlemi ifade eder. Ayrıca ***ε_i***, ***i***nci gözlemin regresyon doğrusuna olan uzaklığını ifade etmekte olup ortalaması 0 ve varyansı ***σ²*** olan bir normal dağılış gösterir

Çoklu doğrusal regresyonda iki veya daha çok sayıda bağımsız değişken veya bağımsız değişken fonksiyonu bulunur. Örneğin, önce verilmiş olan regresyon modeli yeni bir terim *x_i²* eklenerek değiştirilirse; şu anakütle çoklu doğrusal regresyon modeli ortaya çıkar:

$$\text{parabol: } y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i, i = 1, m$$

Denklemin sağ tarafındaki bağımsız değişken için bir ikinci derece (kuadratik) ifade bulunmasına rağmen bu model hala doğrusal regresyon modelidir; çünkü üç tane parametre, yani ***β₀***, ***β₁*** ve ***β₂*** ile çoklu değişkenli doğrusaldır

Daha genel çoklu doğrusal regresyon modelinde *p* tane bağımsız değişken olduğu varsayılır ve anakütle modeli şöyle ifade edilir:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \epsilon_i,$$

veya biraz daha kısa olarak

$$y_i = \beta_0 + \sum_{j=1}^p X_{ij} \beta_j + \epsilon_i$$

İki Değişkenli regresyon katsayı kestirimleri

Anakütleden bir rastgele örneklem elde edilirse, b***β_i*** anakütle parametreleri için örneklem tahminler***β̂_i*** bulunur ve şu örneklem doğrusal regresyon denklemi elde edilir:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + e_i$$

Burada ***e_i*** terimi örneklemden elde edilen *artık* olur; ve

$$e_i = y_i - \hat{y}_i$$

olur.

Bu örneklem regresyon denkleminin kestirimini elde etmenin yöntemi **adi en küçük kareler (Ordinary least squares)** olarak adlandırılır^[1] Bu yönteme göre parametre kestirimleri artıkların kare toplamının minimum (en küçük) değerini bulmakla elde edilirArtıklar minimum toplamıAKT şudur:

$$AKT = \sum_{i=1}^N e_i^2$$

Bu fonksiyonun minimum değerini bulmak her parametre ile birinci kısmal türevi alınarak sıfıra eşitlenir. Böylece her kısmal türeve göre bir denklem elde edilir. Örneğin iki parametrelili doğrusal regresyon için iki değişkenli iki denklem elde edilir. Bu simultane denklem sistemine normal denklemler adı verilir ve bu simultane denklemler sistemi birlikte çözülüp her anakütle parametresi için bir kestrim değeri bulunur

İki parametrelili basit doğrusal regresyon için iki anakütle parametre kestirimleri olan ***β̂₀***, ***β̂₁*** şu normal denklemlerin birlikte çözülmesi ile elde edilir:

$$\begin{aligned} m \alpha + \sum x_i \beta &= \sum y_i \\ \sum x_i \alpha + \sum x_i^2 \beta &= \sum x_i y_i \end{aligned}$$

Burada bütün toplamları i=1 den i=n kadardır ve örneklem toplamı, toplam kareleri ve toplam karşı çarpımı ile elde edilirler

Bu iki değişkenli iki doğrusal denklem için çeşitli çözüm yolları vardırBurada Cramer'in kuralı kullanılıp şu ifade elde edilir:

$$\begin{aligned} \hat{\beta} &= \frac{m \sum x_i y_i - \sum x_i \sum y_i}{\Delta} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \\ \hat{\alpha} &= \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{\Delta} = \bar{y} - \bar{x} \hat{\beta} \end{aligned}$$

burada

$$\Delta = m \sum x_i^2 - \left(\sum x_i \right)^2$$

Böylece iki parametrelili doğrusal regresyon için en küçük kareler parametre tahminleri için formüller şöyledir:

$$\widehat{\beta}_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \text{ a}$$

ve

$$\hat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x}$$

Burada \bar{x} x gözlem değerlerinin ortalaması ve \bar{y} ise y gözlem değerlerinin ortalamasıdır

Çok değişkenli regresyon katsayı kestirimleri

Çok değişkenli regresyon için regresyon katsayısı kestirimi de iki değişkenli regresyon hâli gibi *en küçük kareler* yöntemi ile yapılır.^[1] Bu demektir ki kestirmede ortaya çıkacak artıkların karelerinin toplamı minimize edilecektir. Artıklar r_i olur ve minimize edilecek objektif fonksiyon şöyle ifade edilir:

$$S = \sum_{i=1}^m r_i^2,$$

Her bir artık, gözlemi yapılan bir değer ile kestirim modelini kullanarak elde edilen bir kestirim değeri arasındaki farktır:

$$r_i = y_i - \sum_{j=1}^n X_{ij} \hat{\beta}_j$$

En küçük kareler yöntemine göre S , toplam artıklar karesi, minimize edilecektir. Belirli koşullar geçerli ise, elde edilecek parametre kestirimleri (Gauss-Markov teoremine göre) en küçük varyans gösteren kestirim ve hattamaksimum değışirliliközelliliği gösteren dogrusal kestirim olabilirler

Anakütle regresyon katsayılarınınkestirim regresyon katsayıları için bir $\hat{\beta}$ (şapka) notasyonu kullanılır: yani $\hat{\beta}_i$. Genellikle çok değişkenli regresyon için normal denklemlerşöyle ifade edilir

$$\sum_{i=1}^N \sum_{p=1}^p X_{ij} X_{ik} \hat{\beta}_k = \sum_{i=1}^N X_{ij} y_i, \quad j = 1, p$$

Matris notasyonu kullanılırsa, normal denklemler şöyle yazılabilir:

$$(\mathbf{X}^T \mathbf{X}) \hat{\beta} = \mathbf{X}^T \mathbf{y}$$

Eğer $\mathbf{X}^T \mathbf{X}$ matrisi singular değilse

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y},$$

Bu doğru uydurma tahminleridir

Hatalar varyansı ve toplam kareler

Anakütle hatalarının normal dağılım gösterdiğine dair bir diğer varsayımı da kullanarak incelemeci hatalar varyansı ve toplam kareler değerlerini bulur ve bunları kullanarak tahmin edilen denklem ve parametreler üzerinde çıkarımsal istatistik sonuçlara varabilir^[1]

Anakütle hata teriminin sabit bir varyansı bulunduğu varsayımına göre hatalar varyansı kestirimi şöyle bulunur:

$$\hat{\sigma}_\epsilon = \sqrt{\frac{SSE}{N-2}}$$

Bu ifadeye regresyon için kare kök ortalama hata karesi adı verilir. Parametre kestirimleri için standart hata şöyle bulunur:

$$\hat{\sigma}_{\beta_0} = \hat{\sigma}_\epsilon \sqrt{\frac{1}{N} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2}}$$

$$\hat{\sigma}_{\beta_1} = \hat{\sigma}_\epsilon \sqrt{\frac{1}{\sum (x_i - \bar{x})^2}}$$

Örneklem veri serisinin değişebilirliği değişik toplam kareler suretiyle ifade edilebilirler

- **Tüm toplam kareler** (örneklem varyansına orantılı olur):

$$SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2.$$

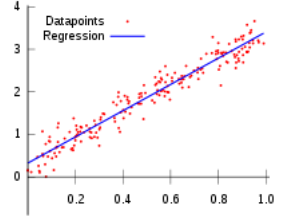
- **Regresyon toplam kareler** Bazan *açıklanan toplam kareler* diye anılır.

$$SS_{\text{reg}} = \sum_i (f_i - \bar{f})^2$$

- **Toplam hatalar karesi** *Artıklar toplam karesi* olarak da isimlendirilir

$$SS_{\text{err}} = \sum_i (y_i - f_i)^2$$

Bir regresyon modeli için parametreler, en küçük kareler yöntemi kullanılarak, tahmin edildikten sonra regresyon kontrol etme yöntemleri (yani çıkarımsal kontrol) uygulamak gerekir. Bunlar bulunan tahmin edilmiş genel doğrusal denklemin örneklem uyum iyiliği sınaması ve kestirimci regresyon katsayılarının tek tek istatistiksel anlamlılığını sınamak veya güvenlik aralığı sağlamak olarak özetlenir



Bir veri serisi için kırmızı noktalar) doğrusal regresyon.

Kestirim denklemin genel uyum iyiliğinin çıkarımsal kontrolü

Tahmin edilen doğrusal regresyon denkleminin genel olarak örnekleme uygun olup olmadığı için uygulanan çıkarımsal istatistik işlemleri iki türde olabilir^[1]

- **belirleme katsayısı** yani **R-kare** değerinin bulunması;
- regresyon denklemine F-sınaması uygulaması.

Bu iki çıkarımsal kontrol aynı konunun çıkarımsal kontrolü için kullanılır: tahmin regresyon denkleminin tümünün uygun olup olmadığı. Bu nedenle iki yöntemden tek birini kullanmak yeterlidir. Buna rağmen birçok pratik analizde her iki çıkarımsal analiz de kullanılmaktadır; çünkü ikisinde hesaplanması için nerede ise aynı yöntem kullanılır ve eğer istatistik veya ekonometrik kompüter paketi kullanılırsa her iki yöntem için gerekli sonuçlar birlikte verilir

Belirleme katsayısı yani R-kare (**R²**) değeri

Belirleme katsayısı yani (**R²**)) **R-kare**) için en genel tanımlama förmülü sudur^[1]

$$R^2 \equiv 1 - \frac{SS_{err}}{SS_{tot}}.$$

yani (*1-toplam hatalar kaesi*) ile (*tüm toplam kareler*) oranı; veya (*1-Artıklar toplam kaesi*) ile (*tüm toplam kareler*) oranı.

Çıkarımsal analizde R-kare değeri bulunur ve bulunan değer doğru hesaplanmışsa 0 ile 1 arasında olmalıdırYapılan bu analiz çeşitli hallerde açıklanabilir:

- Eğer sadece iki değişkenli (Y ve X0) regresyon analizi yapılıyorsa, geometrik olarak X-Y düzeyinde doğrusal regresyon tahmini bu düzeyde olan noktalara en-küçük-kareler prensibine göre en iyi uyan bir doğru uydurmaktırEğer bütün noktalar bu tahmin edilen doğru üzerinde ise tahmin edilen doğru uyumu bu veri noktaları için hiç hatasızdır ve veri noktaları doğrudan hiç ayrılık göstermez. Hat varyasyonu bu halde sıfır olurBu veri noktaları tahmin edilen doğrudan uzaklaştıkça hatalar varyasyonu çoğalmaktadır
- Üç değişkenli (Y ve X1 ve X2 değişkenli) çoklu regresyonda ise tahmin edilen bir üç boyutlu düzeydeEğer bu düzey veri noktalarına tam olarak uyarsa bütün veri noktaları tahmin edilen düzey üstüne düşerler Veri noktaları tahmin edilen düzey etrafında dağılmaya başlarlarsa, hatalar varyasyonu artmaya başlar
- Üç değişkenden daha çoklu değişkenli regresyon analizi ise grafik olarak verilemez; çok boyutlu uzayı gösterir ve tahmin edilen regresyon katsayıları ise bu çok boyutlu uzayda t *hiper-düzey* ortaya çıkarır ve bu*hiper-düzeyden* ayrılmalara hata varyasyonudur

Genel olarak:

- Eğer **R²** değeri sıfıra *yakınsa*, uyum iyiliği uygun olmadığı kabul edilir Bu sonuç ortaya çıkarsa toplanan verilere kullanılan modelin uygun olmadığı sonucu çıkarılır ve bu uygunsuzluk modelinin değiştirilmesini gerektirirBu demektir ki model ile açıklanan varyasyon tüm varyasyonu %0ini açıklamakta ve geometrik olarak örneklem verileri regresyon ile elde edilen hiperdüzeyin etrafına çok dağılmış olarak bulunmaktadırBu çıkarıma varılırsa bu basamağa kadar yapılmış olan analizin bir kenara bırakılması ve diğer bir modelin bulunup kullanılması gerekir
- Eğer **R²** değeri bire *yakınsa*, uyum iyiliği uygun olduğu sonucu çıkarılır Geometrik olarak çok değişken boyutlu uzayda, örneklem veri noktalarının hepsi uzayda dağılma göstermeyip hemen hepsi regresyon ile elde edilen hiperdüzey üzerinde bulunmaktadırBu halde çıkarımsal kontrol analizinin diğer basamağına geçilir

F-testi

Regresyon denklem uyum iyiliği için F-testi için sıfır hipotez (H₀) anakütle model katsayılarının hepsinin değerinin 0 (sıfır)a eşit olduğudur^[1] Yani sıfır hipotez gerçekse hesap ile elde edilen katsayı kestirimlerinin hepsi anakütle için 0dir yani hiçbir etki ve bağlantı anlamı vermez. Bu basamağa kadar yapılmış olan analizin bir kenara bırakılması ve diğer bir modelin bulunup kullanılması gerekir. Eğer sıfır hipotez reddedilirse regresyon kestirimleri genellikle uygundur ve cıkarımsal kontrol analizine devam edilir

Kestirimi yapılan her tek regresyon parametresinin değerinin çıkarımsal kontrolü

Yani bu tahmin edilmiş standart hataları kullanarak her tek anakütle regresyon parametresi hakkında*kestirim güvenlik aralıkları*kurabilir ve *hipotez sınamaları*yapılabilir^[1]

İnterpolasyon ve ekstrapolasyon

Regresyon modelleri kullanarak kestirim yapılmak istenirse, bağımsız değişken olan *x* değişken veri değerleri verilirse bağımlı değişken *y* için kestirim değerleri (**f_i**) tahmin etmek için kullanabilirler

Eğer bu kestirim, modeli kurmak için kullanılan *x* değişken değerleri için yapılıyorsa, bu işlem interpolasyon olarak adlandırılır. Eğer kestirim modeli kurmak için kullanılan bağımsız değişken değerlerinin dışındaki değerler ile yapılırsaekstrapolasyon olarak adlandırılır ve ekstrapolasyon çok daha yanlış olabilir

Diğer yaklaşımlar

Ağırlıklı en küçük kareler yöntemi

En küçük kareler yöntemi kullanılırken yapılan temel varsayımlarından biri hata terimi varyanslarının birbirine eşit olduğudur. Eğer gözlem hataları birbirine eşit olmayan hata terimi varyansları gösteriyorlarsa, en küçük kareler yönteminin bu önemli varsayımı ihlal edilmiş olur ve *en küçük kareler* yöntemi ile elde edilen regresyon kestirimleri anlamlı olmayabilir. Bu sorunu çözmek için her gözleme ayrı ağırlık vererek en küçük kareler yöntemi uygulamak imkânı vardır ve bu genelleştirmeye ağırlıklı en küçük kareler adı verilir. Ağırlıklı En Küçük Kareler Yöntemi,değişkenlere ağırlık vererek veya değişkenlerin önem derecesini değiştirerek uygulanan en küçük kareler yöntemidir

Değişkenlerde-hatalar modeli

En küçük kareler yöntemi kullanılırken yapılan temel varsayımlarından biri de gözlem hatalarının yalnızca bağımlı değişkenlerde yapıldığı ve bağımsız değişkende gözlem hatası bulunmadığıdır. Eğer bağımlı değişkende hata bulunduğu ve bunun elde edilen regresyon kestirim değerlerine çok etki yaptığı bilinirse, degiskenlerde-hatalarmodeli veya diğer ismi ile total en kucuk kareler modeli ve bu modelle ilişkili kestirim yöntemleri kullanılabilir

Genelleştirilmiş doğrusal model

Eğer anakütle regresyon modeli içindeki hatalar için olasılık dağılım fonksiyonu bir normal dağılım göstermiyorsa genelleştirilmiş doğrusal model kullanılabilir. Bu halde hataların olasılık dağılım fonksiyonu ussel dağılım, gamma dağılımı, ters Gauss tipi dağılım, Poisson dağılımı, binom dağılım, multinom dağılım vb. olabilir

Güçlü regresyon

Ayrık bağımlı değişken

Doğrusal olmayan regresyon

Eğer değişkenler hakkında kurulan teori dolayısıyla anakütle modeli parametreleri ile doğrusal değilse, kestirim yapılırken toplam kareleri tekrarlama usulu kullanarak minimize edilmesi gerekir. Bu kullanılan tekrarlama yöntemi birçok karışık sorunlar ortaya çıkarır. Bunların özet olarak incelenmesi için [doğrusal olmayan regresyon](#)[#Doğrusal olmayan ve doğrusal en küçük kareler arasındaki farklar](#) maddesine bakınız.^[1]

İçsel kaynaklar

- Parçalar için regresyon
- Güvenlik aralığı
- Güvenlik yöreni
- Ekstrapolasyon
- Krigleme
- Tahmin etme
- Tahmin aralığı
- İstatistik
- Trend kestirimi
- Güçlü regresyon
- Çokdeğişirli normal dağılım

Kaynakça

- ↑ ***a b c d e f g h i j k*** Gujarati, Damodar (çev Ümit Şenesen, Gülâş Günlük Şenesen) (2008)*Temel Ekonometri* Literatür Yayınları ISBN 975-7860-99-9
- ↑ A.M. Legendre (1805),*Nouvelles méthodes pour la détermination des orbites des comètes*"Sur la Méthode des moindres quarrés" bir ek bölümde bulunur
- ↑ C.F. Gauss (1809), *Theoria Motus Corporum Coelestium in Sectionibus Conicis Şölem Ambientum*
- ↑ C.F. Gauss (1821/1823).*Theoria combinationis observationum erroribus minimis obnoxiae*
- ↑ Francis Galton (1877), "Typical laws of heredity", *Nature* 15, 492-495, 512-514, 532-533(*Galton burada bezelyelerle yaptığı kalıtım deneyi sonucunda*[eversion terimi kullanır](#))
- ↑ Francis Galton (1885) Presidential address, Section H, Anthropology(*Burada insanların boyları üzerinde yaptığı araştırma sonucu için "regression" terimi kullanılır*)
- ↑ G. Udney Yule (1897) "On the Theory of Correlation", *J. Royal Statist. Soc.*, 1897, p. 812-54.
- ↑ Karl Pearson, G.U. Yule, Norman Blanchard, and Alice Lee (1903). "The Law of Ancestral Heredity"*Biometrika*
- ↑ R.A. Fisher (1922), "The goodness of fit of regression formulae, and the distribution of regression coefficients", *J. Royal Statist. Soc.*, 85, 597-612
- ↑ R.A. Fisher (1925)*Statistical Methods for Research Workers*

Bibliyografya

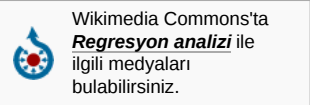
- Gujarati, Damodar (çev Ümit Şenesen, Gülâş Günlük Şenesen) (2008)*Temel Ekonometri* Literatür Yayınları ISBN 975-7860-99-9
- Audi, R., Ed. (1996). "Curve fitting problem,"*The Cambridge Dictionary of Philosophy*Cambridge, Cambridge University Press. pp. 172–173(İngilizce)
- William H. Kruskaland Judith M. Tanur, ed. (1978), "Linear Hypotheses,"*International Encyclopedia of Statistics*Free Press, cilt 1(İngilizce)

Evan J. Williams, "I. Regression," say. 523-41.
Julian C. Stanley, "II. Analysis of Variance," pp. 541-554.

- Lindley, D.V. (1987). "Regression and correlation analysis,"*New Palgrave: A Dictionary of Economics* Cilt. 4, say. 120-23. (İngilizce)
- Birkes, David and Adolah Dodge, *Alternative Methods of Regression* ISBN 0-471-56881-3(İngilizce)
- Chatfield, C. (1993) "Calculating Interval Forecasts,"*Journal of Business and Economic Statistics***11**. pp. 121–135. (İngilizce)
- Diğer, F. "Regresyon Kelimesinin Tarihi".^[1]
- Draper, N.R. ve Smith, H. (1998)*Applied Regression Analysis*Wiley Series in Probability and Statistics(İngilizce)
- Fox, J. (1997). *Applied Regression Analysis, Linear Models and Related Methods*Sage (İngilizce)
- Hardle, W., *Applied Nonparametric Regression*(1990), ISBN 0-521-42950-1(İngilizce)
- Meade, N. ve T. Islam (1995) "Prediction Intervals for Growth Curve Forecasts,"*Journal of Forecasting*,**14**, say. 413-430. (İngilizce)
- Munro, Barbara Hazard (2005) "Statistical Methods for Health Care Research 5th ed." Lippincott Williams & Wilkins(İngilizce)
- Sykes, A.O. "An Introduction to Regression Analysis"(Innaugural Coase Lecture)
- Kotsiantis, S., Kanellopoulos, D. ve Pintelas, P(2006) "Local Additive Regression of Decision Stumps"*Lecture Notes in Artificial Intelligence* Springer-Verlag, Vol. 3955, SETN 2006, say. 148 – 157 (İngilizce)
- Kotsiantis, S. ve Pintelas, P (2005) "Selective Averaging of Regression Models"*Annals of Mathematics, Computing & Teleinformatics*, Cilt 1, No 3, say 66-75 (İngilizce)

Dış bağlantılar

- Regresyon analizi
- Curvefit: A complete guide to nonlinear regression Online ders kitabı
- RM4E ile regresyonun kolaylaştırılması
- Doğrusal model üzerine yorumlar- Bill Venables tarafından bazı doğrusal regresyon modelleri üzerinde birkaç yorum.
- Mazoo's Learning Blog- Doğrusal regresyon için örnek. Doğrusal regresyon denkleminin, varyansların, standart hataların, korelasyon katsayısının, belirleme katsayısının ve güvenlik aralıklarının nasıl bulunduğunu göstermektedir
- Zayıf korelasyon bağlantılı verilerin regresyonu- Y -aralığı X-aralığından çok daha küçük olursa nasıl regresyon hataları ortaya çıkabilir
- xuru.org Online regresyon avadanlığı
- Matlab SURrogate MOdeling Toolbox - SUMO Toolbox - Aktif öğrenme + Model seçimi + Çoklu model regresyonu için Matlab yazılımı



 G • T • D	İstatistik		
Betimsel istatistik	Sürekli veriler	Merkezi konum	Ortalama (Aritmetik, Geometrik, Harmonik) • Medyan • Mod
		Yayılma	Açıklık • Standart sapma • Varyasyon katsayısı • Çeyrekler açıklığı • Kesirlilikler (kantil) (Dörttebirlik, Ondabirlik, Yüzdebirlik)
		Dağılım şekli	Varyans • Çarpıklık • Basıklık • Moment (matematik)
	İstatistiksel tablolar	Sıklık dağılımı	Çoklu sayılı özetleme tabloları • İlişki tablosu • Çoklu-yönlü sınıflandırma tabloları
Veri toplama	İstatistiksel grafikler	Dairesel grafik • Çubuk grafiği • Kutu grafiği • Dal-yaprak grafikleri • Kontrol diyagramı • Histogram • Sıklık çizelgesi • Q-Q grafiği • Serpilme diyagramı	
	Örnek tasarımı	Anakütle • Basit rassal örnekleme • Örüntülü örnekleme • Tabakalı örnekleme • Küme örnekleme • Çok aşamalı örnekleme •	
	Deneyisel tasarım	Anakütle • İstatistiksel deneyel tasarım tipleri • Deneyel hata • Yineleme • Bloklama • Duyarlılık ve belirleme	
Çıkarımsal istatistik ve istatistiksel kestirim ve testler	Örneklem kavramları	Örneklem büyüklüğü • Sınama gücü • Etki büyüklüğü • Örnekleme dağılımı • Standart hata	
	Çıkarımsal analiz tipleri	Kestirim • Parametrik çıkarımsal analiz • Parametrik olmayan çıkarımsal analiz • Bayesci çıkarımsal analiz • Meta-analiz	
	Çıkarımsal kestirim	Genel kestirim kavramları	Momentler yöntemi • Maksimum olabilirlik • Bayes-tipi kestirimci •

			Minimum uzaklık • Maksimum aralık verme
		Tekdeğişkenli kestirim	Kestirim • Güven aralığı • İnanılır aralık
	Hipotez testi	İstatistiksel test ana kavramları	Sıfır hipotez • I.Tür ve II.Tür hata • Anlamlılık seviyesi • p-değeri
		Basit tek-değişkenli ve iki-değişkenli parametrik hipotez testi	μ için testi • π için test • μ_1 - μ_2 için test • π_1 - π_2 için test • σ_1/σ_2 için test
Tek-değişkenli ve iki-değişkenli parametrik olmayan test analizi		Medyan testi • Ki-kare testi • Pearson ki-kare testi •Phi katsayısı • Wald testi • Mann-Whitney U testi • Wilcoxon'ın işaretli sıralama testi	
Korelasyon ve Regresyon analizi	Korelasyon	Pearson çarpım-moment korelasyonu • Sıralama korelasyonu (Spearman'ın rho • Kendall'in tau)	
	Doğrusal regresyon	Regresyon analizi • Doğrusal model • Genel doğrusal model • Genelleştirilmiş doğrusal model	
	Doğrusal olmayan regresyon	Parametrik olmayan • Yarıparametrik • Logistik	
	Varyans analizi	Tek-yönlü varyans analizi • Kovaryans analizi • Bloklü tek-yönlü varyans analizi • Etki karışımı değişkeni	
Çokdeğişkenli istatistik	Çokdeğişkenli regresyon • temel bileşenler • Faktör analizi • Kanonik korelesyon • Uygunluk analizi • Kümeleme analizi		
Zaman serileri analizi	Yapısal model tanımlanması	Zaman serisi yapısal model ögeleri • Zaman serisi ögeleri saptanması • Zaman grafiği • Korrelogram	
	Zaman serileri kestirim teknik ve modelleri	Dekompozisyon • Trend uygulama kestirimi • Üssel düzgünleştirme • ARIMA modelleri • Box–Jenkins • Spektral yoğunluk kestirimi	
	Kestirim değerlendirmesi	Zaman seri kestirim değerlendirmesi	
Sağkalım analizi	Sağkalım fonksiyonu • Kaplan–Meier • Log-sıra testi • Başarısızlık oranı • orantılı tehlikeler modeli		
Kategori • Outline • Endeks			
G • T • D			
Genel	Halk sağlığı		
	Oksoloji • Biological hazard • Chief Medical Officer • Cultural competence • Deviance • Çevre sağlığı • Euthenics • Genomics • Global Health economics • Health literacy • Sağlık politikaları • (Health system • Healthcare reform •Public health law) •Maternal health •Tıbbi antropoloji •Medical sociology •Mental health •Pharmaceutical policy •Public health • Sosyal psikoloji • Sociology of health and illness • Tropical disease		
Koruyucu sağlık hizmetleri	Behavior change • Aile planlaması • Health promotion • Human nutrition • Hijyen • (El yıkama • Infection control • Oral hygiene) • İş güvenliği • Injury prevention • Medicine • Nursing) • Patient safety • (Organization) • Pharmacovigilance • Safe sex • (Cinsel yolla bulaşan hastalıkların önlenmesi) • (Fecal-oral transmission • Open defecation • Vector control • Waterborne diseases) • Smoking cessation • Aşılama		
Population health	Biyoistatistik • Child mortality • Community health • Epidemiyoloji • Global health • Health impact assessment • Health system • Bebelik • Open-source healthcare software • Public health informatics • Social determinants of health • (Health equity • Race and health) • Social epidemiology		
Biyolojik ve epidemiyolojik istatistikler	Hipotez testi • case-control study • randomized controlled trials • Klinik epidemiyoloji • ROC eğrisi • Öğrencilere t-testi • Z-test • Varyans • statistical software • (SAS • SPSS • Stata)		
Bulaşıcı ve epidemik hastalıkların önlenmesi	Cinsel yolla bulaşan hastalık • vaccine regulation • Aşılama • tropical diseases • Enfeksiyon hastalıkları • Epidemiyoloji • surveillance •		
Gıda hijyeni ve sertifikasyon sistemleri	Gıda güvenliği • Gıda katkısı • Gıda işleme • Certified Agricultural • GMP • GSP • GAP • CSP • ISO • HACCP • CCP • Halal (for Islamic Food microbiology • food Chemistry • food technology • Genetik mühendisliği		
Sağlık davranışları bilimleri	Health belief model • rational behavior model • theory of planned behavior • Transtheoretical model • Social cognitive theory • Social S • precede porceed model • Community health • Ecosystem model		
Sağlık ve politika analizi	policy analysis • Insurance Economics • İstatistik • regional regulations • social security system • Muhasebe • Sosyal psikoloji • İletişim		
Organizasyonlar Eğitim ve Geçmiş	Kuruluş ve Örgütler	Europe (Avrupa Hastalık Önleme ve Kontrol Merkezi • Committee on the Environment, Public Health and Food Safety) • India (Ministry of Health and Family Welfare)	
		U.S. (Amerika Birleşik Devletleri Hastalık Kontrol ve Korunma Merkezleri • Center for Minority Health • County of Los Angeles • Public Health – Seattle & King County • Public Health Service) • Globalization and Health • Dünya Sağlık Örgütü • World Health Organization	
	Eğitim	East Asia (Health and Welfare Ministry • Gıda ve İlaç Dairesi • Çevre Koruma Ajansı) • Bachelor of Science in Public Health • Master of Public Health • Doctor of Public Health • European Programme for Intervention Epidemiology Training (EPIET) • Health education • Professional Further Education in Clinical Pharmacy and Public Health	
		Geçmiş Sara Josephine Baker • Samuel Jay Crumbine • Carl Rogers Darnall • Joseph Lister • Margaret Sanger • John Snow • Hastalık yapıcı mikrop teorisi • Social hygiene movement	

"https://tr.wikipedia.org/w/index.php?title=Regesyon_analizi&oldid=19535349" adresinden alındı.

Bu sayfa son olarak 16 Şubat 2018 tarihinde ve 20.09 saatinde düzenlenmiştir

Metin Creative Commons Atıf-BenzerPaylaşım Lisansaltındadır; ek koşullar uygulanabilir Bu siteyi kullanarak, Kullanım Şartlarınıve Gizlilik Politikasınıkabul etmiş olursunuz. Vikipedi® (ve Wikipedia®) kâr amacı gütmeyen kuruluş olanWikimedia Foundation, Inc.tescilli markasıdır