

MARMARA ÜNİVERSİTESİ  
FEN-EDEBİYAT FAKÜLTESİ  
İSTATİSTİK BÖLÜMÜ

---

Assiye Sevde ÖZATA 121518964  
Ramazan YALÇIN 121519821



# İçindekiler

- Panel Veri Hakkında Bilgi
- Günümüzde Panel Veri Nerede?
- Panel Verinin Amacı
- Veri Hakkında Bilgi
- Veri Hakkında Özet Bilgi
- Korelasyon Matrisi
- Görselleştirme
- Regresyon Modeli
- Havuzlanmış En Küçük Kareler
- Sabit Etkili Model
  - En Küçük Kareler Kukla Değişkeni Modeli (LSDV)
  - Within Model Kullanılarak Sabit Etkiler
- Bireysel Etkiler İçin F Testi
- Rastgele Etkili Model
- Hausman Testi
- Değişen Varyanslılık
- Box-Cox Dönüşümü
- Sonuç
- Kaynakça

# Panel Veri Hakkında Bilgi

Panel veri, yatay kesit verileri ile zaman serilerini bir araya getirmektedir. Buna göre daha çok aydınlatıcı veri, daha çok değişkenlik, değişkenler arasında daha az doğrusal bağlantı, daha çok serbestlik derecesi ve daha çok etkinlik sağlamaktadır. Zaman boyutuna sahip yatay kesit veriler bir başka ifade ile panel veriler kullanılarak oluşturulan panel veri modelleri yardımıyla veri seti ilişkilerinin tahmin edilmesi yöntemine “panel veri analizi” adı verilmektedir.

Panel veri analizinde genelde, yatay kesit birim sayısının  $N$  dönem sayısından  $T$  fazla ( $N > T$ ) olduğu durumla karşılaşılmaktadır. Genel olarak panel veri modeli;

$$Y_{it} = \alpha + \sum_{k=1}^q \beta_k X_{kit} + u_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T; \quad k = 1, 2, \dots, q$$

şeklinde yazılabilmektedir. Burada,  $Y$  (bağımlı değişken),  $X_k$  (bağımsız değişkenler),  $\alpha$  sabit parametre,  $\beta_k$  eğim parametreleri ve  $u_{it}$  hata terimidir.  $i$  alt indisi yatay kesit birimlerini (birey, firma, şehir, ülke gibi),  $t$  alt indisi ise zamanı (gün, ay, yıl gibi) ifade etmektedir. Hata terimi  $u_{it}$ 'nin ortalamasının sıfır ve sabit varyanslı olduğu varsayılmaktadır. Bu modelde sabit ve eğim parametreleri hem birimlere hem de zamana göre değer almaktadır.



# Günümüzde Panel Veri Nerede?

---

Panel veri analizi, birim-zaman boyutu olmak üzere iki boyutlu verilerin analizidir. Sosyal bilimlerin yanı sıra birçok disiplinde bir araştırma yöntemi olarak kullanımı giderek artmaktadır.

Özellikle ekonomi ve finans alanında yoğun olarak kullanılmaktadır. Bu alanların dışında da sağlık, endüstri, işletme gibi alanlarda da panel veri analizi karşımıza çıkmaktadır.

Panel veri analizi için bazı uygulamalar şunlardır:

- 2010-2020 yılları arasında OECD ülkelerindeki işsizlik oranlarını etkileyen faktörlerin belirlenmesi.
- 2019 yılında aylık bazda hastanelerin doluluk oranları üzerinde etkisi olan göstergelerin incelenmesi.
- 2000-2020 yılları arasında farklı döviz kurlarının (Dolar, Euro vs.) hisse senetlerinin getiri tutarları üzerindeki etkilerinin değerlendirilmesi.

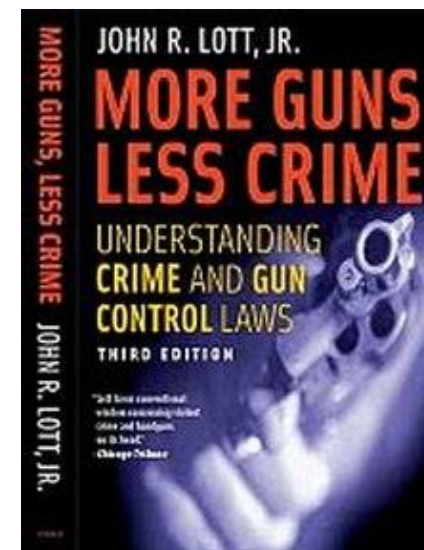
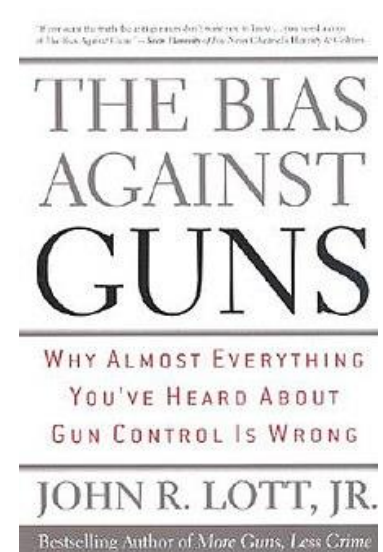
# Panel Verinin Amacı

Panel veri analizinde örneklemdaki farklı zaman noktaları için bireysel gözlemler dikkate alınır ve bu örneklemdaki her bir bireysel veri için çoklu gözlemler oluşturulması sağlanır. Yatay kesit veri birçok birim için sadece bir dönem hakkında bilgi verirken, zaman serisi verisi sadece bir birimin dönemlere göre bilgisini vermektedir. Hem dönemlere hem de birimlere göre bilgiler isteniyorsa, panel veri kullanılması gerekmektedir.

Panel veri, zaman serisi ve yatay kesit verisinden daha karmaşık davranışsal modeller yapmaya ve bunları test etmeye yardımcı olur. Bizim amacımız bu panel veriyi analiz ederek veri setimiz hakkında sonuçlarla yorumlar yapabilmektir.



# John Richard Lott Jr.



# Amerika Birleşik Devletleri Suç Oranını Etkileyen Etmenler: Panel Veri Analizi

Veri, Amerika Birleşik Devletleri'ndeki 50 eyalete ait 1977-1999 yılları arasında;

- **mur**-Cinayet oranı (100.000'de bir olay)
- **rob**-Soygun oranı (100.000' de bir olay)
- **vio**-Şiddet içeren suç oranı (100.000'de bir olay)
- **shall**-Devletin o yıl yürürlükte olan bir taşıma yasası varsa 1, aksi takdirde 0
- **incarc\_rate**-Bir önceki yıl eyaletteki hapsedilme oranı (100.000 kişi başına hükümlü mahkum; önceki yıl için değer)
- **pm1029**-10 ile 29 yaşları arasındaki erkek eyalet nüfusunun yüzdesi
- **pw1064**-Beyaz olan ve 10 ile 64 yaş arasındaki eyalet nüfusunun yüzdesi
- **pb1064**-Siyahi olan ve 10 ile 64 yaş arasındaki eyalet nüfusunun yüzdesi
- **avginc**-Ortalama devlet geliri (1983'te bin dolar olarak)
- **density**-Mil kare başına nüfus, bin olarak
- **pop**-Eyalet nüfusu, milyon
- **year**-Yıl (1977-1999)
- **stateid**-Devlet tanımlayıcısı

değişkenlerinden oluşmaktadır.

Veri 1150 gözlem, 13 değişken, 23 yıl ve 50 eyaletten oluşmaktadır.

# Veri hakkında Özet Bilgi

```
##      stateid      year      vio      mur
## Length:1150      Min.   :1977      Min.   : 47.0      Min.   : 0.200
## Class :character  1st Qu.:1982      1st Qu.: 281.5      1st Qu.: 3.600
## Mode  :character  Median :1988      Median : 437.5      Median : 6.300
##                               Mean  :1988      Mean   : 472.2      Mean   : 6.833
##                               3rd Qu.:1994      3rd Qu.: 635.2      3rd Qu.: 9.600
##                               Max.   :1999      Max.   :1244.3      Max.   :20.300
##      rob      incarc_rate      pb1064      pw1064
## Min.   : 6.4      Min.   : 19.0      Min.   : 0.2482      Min.   :23.52
## 1st Qu.: 70.2      1st Qu.:113.0      1st Qu.: 2.1603      1st Qu.:60.29
## Median :122.4      Median :183.0      Median : 3.9651      Median :65.34
## Mean   :143.7      Mean   :211.5      Mean   : 4.9652      Mean   :63.71
## 3rd Qu.:188.5      3rd Qu.:285.0      3rd Qu.: 6.4426      3rd Qu.:69.29
## Max.   :684.0      Max.   :736.0      Max.   :26.9796      Max.   :76.53
##      pm1029      pop      avginc      density
## Min.   :12.51      Min.   : 0.4027      Min.   : 8.555      Min.   :0.0007071
## 1st Qu.:14.66      1st Qu.: 1.2424      1st Qu.:11.897      1st Qu.:0.0315201
## Median :15.89      Median : 3.2889      Median :13.316      Median :0.0795801
## Mean   :16.09      Mean   : 4.9006      Mean   :13.623      Mean   :0.1636141
## 3rd Qu.:17.53      3rd Qu.: 5.7378      3rd Qu.:15.162      3rd Qu.:0.1682232
## Max.   :22.35      Max.   :33.1451      Max.   :23.647      Max.   :1.0976428
##      shall
## Min.   :0.0000
## 1st Qu.:0.0000
## Median :0.0000
## Mean   :0.2478
## 3rd Qu.:0.0000
## Max.   :1.0000
```

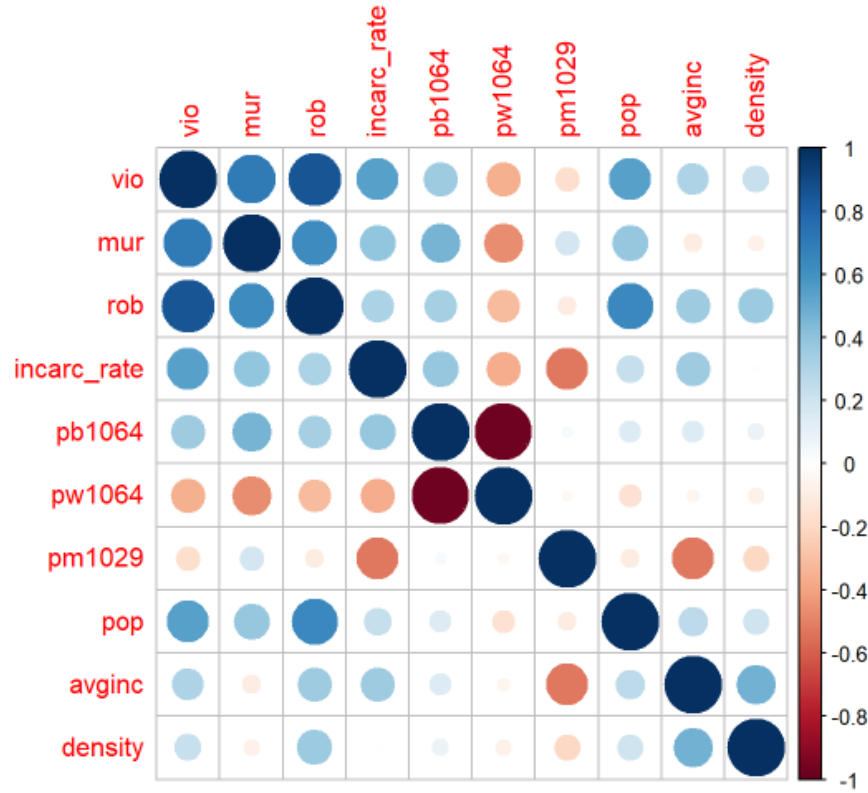
Veri setimizin özet istatistiği yukarıdaki gibidir.  
Verimizde eksik değer bulunmamaktadır.

Panel verilerle çalışıldığında, her bir birim tüm zamanlar boyunca gözlenmişse “dengeli panel”; bazı birimler için bazı zamanlar kayıpsa “dengesiz panel” olarak adlandırılmaktadır.

```
## # A tibble: 50 x 2
##   stateid      n
##   <chr>    <int>
## 1 AK         23
## 2 AL         23
## 3 AR         23
## 4 AZ         23
## 5 CA         23
## 6 CO         23
## 7 CT         23
## 8 DE         23
## 9 FL         23
## 10 GA        23
## # ... with 40 more rows
```

Verimiz dengeli panel veri yapısına sahiptir. ABD’de her bir eyaleti her yıl bir kez olmak üzere 23 kez gözlemliyoruz.





##	vio	mur	rob	incarc_rate	pb1064
## vio	1.0000000	0.70744229	0.8528373	0.541705635	0.35971990
## mur	0.7074423	1.0000000	0.6264604	0.391961033	0.46200479
## rob	0.8528373	0.62646045	1.0000000	0.318500077	0.33052950
## incarc_rate	0.5417056	0.39196103	0.3185001	1.000000000	0.38678061
## pb1064	0.3597199	0.46200479	0.3305295	0.386780611	1.000000000
## pw1064	-0.3504384	-0.46250375	-0.3187807	-0.361497706	-0.97466958
## pm1029	-0.1782223	0.18166544	-0.1037137	-0.523348475	0.03865878
## pop	0.5410316	0.38495083	0.6493391	0.236904915	0.14213137
## avginc	0.3010535	-0.10121138	0.3592404	0.351626351	0.14714389
## density	0.2215428	-0.07574882	0.3673820	-0.006658938	0.08747197
##	pw1064	pm1029	pop	avginc	density
## vio	-0.35043839	-0.17822233	0.5410316	0.30105350	0.221542840
## mur	-0.46250375	0.18166544	0.3849508	-0.10121138	-0.075748817
## rob	-0.31878075	-0.10371372	0.6493391	0.35924041	0.367382011
## incarc_rate	-0.36149771	-0.52334847	0.2369049	0.35162635	-0.006658938
## pb1064	-0.97466958	0.03865878	0.1421314	0.14714389	0.087471975
## pw1064	1.00000000	-0.03766792	-0.1546072	-0.05046704	-0.075059488
## pm1029	-0.03766792	1.00000000	-0.1044715	-0.52686139	-0.203811149
## pop	-0.15460720	-0.10447148	1.00000000	0.26322274	0.201468013
## avginc	-0.05046704	-0.52686139	0.2632227	1.00000000	0.476453893
## density	-0.07505949	-0.20381115	0.2014680	0.47645389	1.000000000

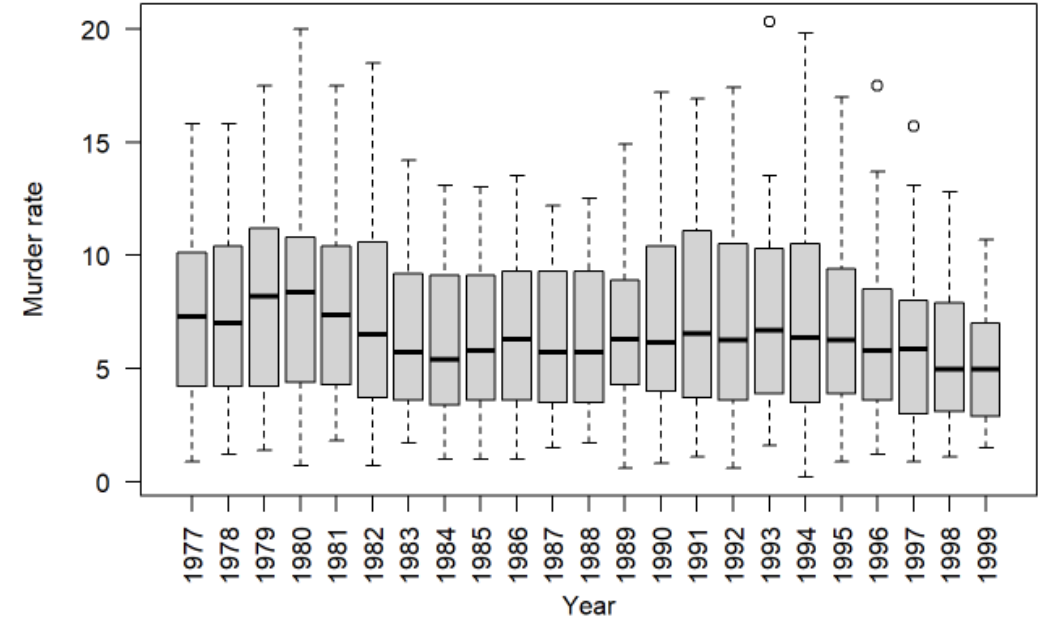
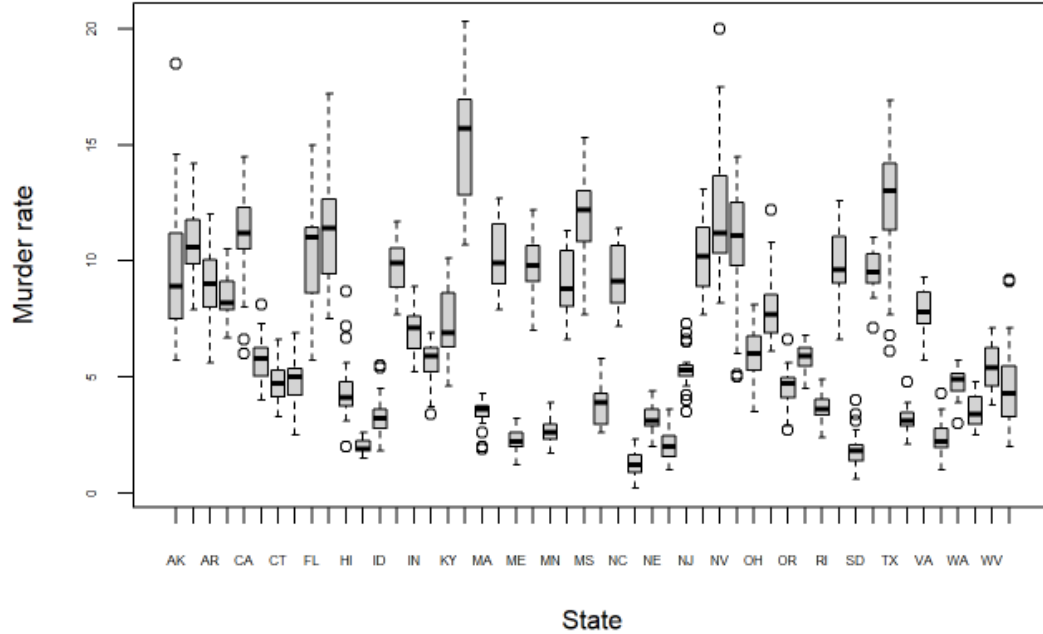
# Korelasyon Matrisi

Cinayet oranına (mur) bakıldığında;

- Şiddet içeren suç oranı (vio) ile yüksek bir korelasyona sahip olduğunu görmekteyiz.
- Mil kare başına nüfus (density) ile negatif ilişkili ve çok düşük korelasyona sahiptir.

# Görselleştirme

Yalnızca bir yıllık verileri gözlemleyecek olsaydık, sonucumuzdaki tek değişiklik eyaletler arasındaki farklılıklardan olacaktı. Ancak verimiz panel yapıya sahip olduğu için cinayet oranlarında da zaman içinde farklılıklar görülmektedir.



# Regresyon Modeli

```
##
## Call:
## lm(formula = mur ~ vio + rob + incarc_rate + pop + pb1064 + pw1064 +
##     pm1029 + avginc + density, data = weapons)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.6275 -1.3127 -0.0847  1.1312 10.2163
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -8.9426790   2.5011308  -3.575 0.000364 ***
## vio          0.0061213   0.0005510  11.109 < 2e-16 ***
## rob          0.0112355   0.0013078   8.591 < 2e-16 ***
## incarc_rate  0.0063725   0.0007645   8.335 < 2e-16 ***
## pop          0.0077206   0.0148654   0.519 0.603604
## pb1064       0.3703701   0.0776226   4.771 2.07e-06 ***
## pw1064       0.1114488   0.0383018   2.910 0.003687 **
## pm1029       0.4634724   0.0503669   9.202 < 2e-16 ***
## avginc      -0.4548707   0.0389082 -11.691 < 2e-16 ***
## density     -1.9222298   0.3169250  -6.065 1.79e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.965 on 1140 degrees of freedom
## Multiple R-squared:  0.7283, Adjusted R-squared:  0.7261
## F-statistic: 339.5 on 9 and 1140 DF, p-value: < 2.2e-16
```

1977-1999 yılları arasını kapsayan panel veride EKK yöntemi ile tahminleme yapılmıştır. Analiz sonuçlarına göre bu dönemde eyaletlerde cinayet oranındaki varyasyonu açıklaması bakımından eyalet nüfusu değişkeni hariç diğer değişkenlerin istatistiksel olarak anlamlı bir etkisi olduğu gözlemlenmektedir.

Şiddet suç oranı, soygun oranı, hapsedilme oranı, eyalet nüfusu, 10 ile 64 yaş arasındaki beyaz olan eyalet nüfusunun yüzdesi, 10 ile 64 yaş arasındaki siyah eyalet nüfusunun yüzdesi ve 10 ile 29 yaşları arasındaki erkek değişkenlerinin cinayet oranı üzerinde pozitif etkisi olduğu gözlemlenmiştir.

# Havuzlanmış En Küçük Kareler

```
## Pooling Model
##
## Call:
## plm(formula = mur ~ vio + rob + incarc_rate + pop + pb1064 +
##      pw1064 + pm1029 + avginc + density, data = weapons, model = "pooling",
##      index = c("stateid", "year"))
##
## Balanced Panel: n = 50, T = 23, N = 1150
##
## Residuals:
##      Min.      1st Qu.      Median      3rd Qu.      Max.
## -6.627468 -1.312705 -0.084719  1.131169 10.216324
##
## Coefficients:
##              Estimate Std. Error t-value Pr(>|t|)
## (Intercept) -8.94267901  2.50113077  -3.5755 0.0003643 ***
## vio          0.00612128  0.00055102  11.1089 < 2.2e-16 ***
## rob          0.01123553  0.00130782   8.5910 < 2.2e-16 ***
## incarc_rate  0.00637245  0.00076450   8.3354 < 2.2e-16 ***
## pop          0.00772064  0.01486540   0.5194 0.6036040
## pb1064       0.37037007  0.07762261   4.7714 2.066e-06 ***
## pw1064       0.11144876  0.03830184   2.9097 0.0036874 **
## pm1029       0.46347235  0.05036691   9.2019 < 2.2e-16 ***
## avginc      -0.45487068  0.03890818 -11.6909 < 2.2e-16 ***
## density     -1.92222978  0.31692505  -6.0653 1.790e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    16193
## Residual Sum of Squares: 4399.8
## R-Squared:              0.72829
## Adj. R-Squared: 0.72615
## F-statistic: 339.522 on 9 and 1140 DF, p-value: < 2.22e-16
```

Havuzlanmış EKK tahmini, panel veri üzerinde çalışan bir EKK tekniğidir. Panel veri, zaman ve kesit boyutları bilgilerini içerir ancak havuzlanmış EKK panel verilerin bu bilgilerini dikkate almaz. Bu nedenle bireye özgü tüm etkiler tamamen göz ardı edilir.

Analiz sonucunda, eyalet nüfusu değişkeni hariç diğer değişkenlerin bireysel etkileri p değerleri 0.05'in altında olduğu için önemlidir.

$R^2$ , uygunluk ölçüsünün iyi olduğunu gösteren 0.72829'a eşittir. Düzeltilmiş  $R^2$ , modelin güvenilir olduğunu gösteren 0.72615'e eşittir.

**Havuzlanmış EKK'nın p-değeri 0.05'ten az olsa da,  $R^2$  ve Düzeltilmiş  $R^2$  genel doğruluğun iyi olduğunu, modelin kullanıma uygun olduğunu gösterse bile heterojenliği tamamen göz ardı ettiği için bu modele güvenemeyiz.**



# Sabit Etkili Model

## En Küçük Kareler Kukla Değişkeni Modeli (LSDV)

```
## Call:
## lm(formula = mur ~ vio + rob + incarc_rate + pop + pb1064 + pw1064 +
##   pm1029 + avginc + density + factor(stateid) - 1, data = weapons)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.7994 -0.6290 -0.0535  0.6060  6.5970
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## vio              2.678e-03  6.695e-04  4.000 6.75e-05 ***
## rob              1.636e-02  1.596e-03  10.253 < 2e-16 ***
## incarc_rate     -6.966e-03  7.343e-04 -9.485 < 2e-16 ***
## pop             -4.200e-01  6.345e-02 -6.619 5.66e-11 ***
## pb1064           6.391e-01  1.324e-01  4.827 1.59e-06 ***
## pw1064           2.827e-02  3.616e-02  0.782 0.434543
## pm1029           2.552e-01  4.629e-02  5.514 4.38e-08 ***
## avginc           2.519e-01  4.617e-02  5.455 6.06e-08 ***
## density          -8.673e+00  4.439e+00 -1.954 0.051004 .
## factor(stateid)AK -7.527e+00  2.808e+00 -2.681 0.007457 **
## factor(stateid)AL -2.396e+00  2.694e+00 -0.890 0.373923
## factor(stateid)AR -2.648e+00  2.546e+00 -1.040 0.298369
## factor(stateid)AZ -3.508e+00  2.568e+00 -1.366 0.172096
## factor(stateid)CA  4.036e+00  3.209e+00  1.258 0.208800
## factor(stateid)CO -6.165e+00  2.699e+00 -2.284 0.022542 *
## factor(stateid)CT -3.763e+00  4.233e+00 -0.889 0.374150
## factor(stateid)DE -7.818e+00  3.252e+00 -2.404 0.016385 *
## factor(stateid)FL -6.766e-01  2.798e+00 -0.242 0.808983
## factor(stateid)GA -3.152e+00  2.806e+00 -1.123 0.261483
## factor(stateid)HI -2.001e+01  3.763e+00 -5.318 1.27e-07 ***
## factor(stateid)IA -6.788e+00  2.594e+00 -2.616 0.009010 **
## factor(stateid>ID -5.757e+00  2.563e+00 -2.246 0.024912 *
## factor(stateid)IL -3.371e+00  2.880e+00 -1.170 0.242123
## factor(stateid)IN -2.223e+00  2.814e+00 -0.790 0.429581
## factor(stateid)KS -5.682e+00  2.561e+00 -2.219 0.026697 *
## factor(stateid)KY -1.947e+00  2.720e+00 -0.716 0.474317
## factor(stateid)LA -7.511e-01  2.763e+00 -0.272 0.785823
## factor(stateid)MA -3.147e+00  4.620e+00 -0.681 0.495863
## factor(stateid)MD -5.905e+00  3.595e+00 -1.643 0.100721
## factor(stateid)ME -6.525e+00  2.642e+00 -2.470 0.013663 *
## factor(stateid)MI -1.545e+00  2.819e+00 -0.548 0.583895
## factor(stateid)MN -7.791e+00  2.648e+00 -2.943 0.003322 **
## factor(stateid)MO -2.792e+00  2.600e+00 -1.074 0.283228
## factor(stateid)MS -2.781e+00  2.734e+00 -1.017 0.309258
## factor(stateid)MT -6.029e+00  2.563e+00 -2.352 0.018837 *
## factor(stateid)NC -2.634e+00  2.828e+00 -0.931 0.351864
## factor(stateid)ND -8.740e+00  2.549e+00 -3.429 0.000629 ***
## factor(stateid)NE -7.322e+00  2.539e+00 -2.884 0.004006 **
## factor(stateid)NH -6.786e+00  2.810e+00 -2.415 0.015911 *
## factor(stateid)NJ -6.859e-01  5.517e+00 -0.124 0.901086
## factor(stateid)NM -3.804e+00  2.611e+00 -1.457 0.145399
## factor(stateid)NV -4.509e+00  2.620e+00 -1.721 0.085586 .
## factor(stateid)NY -1.832e+00  3.251e+00 -0.564 0.573176
## factor(stateid)OH -1.324e+00  3.026e+00 -0.438 0.661834
## factor(stateid)OK -3.466e+00  2.601e+00 -1.333 0.182936
## factor(stateid)OR -7.099e+00  2.586e+00 -2.745 0.006157 **
## factor(stateid)PA -1.126e+00  3.006e+00 -0.374 0.708182
## factor(stateid)RI -6.632e-01  5.374e+00 -0.123 0.901807
## factor(stateid)SC -5.027e+00  2.861e+00 -1.757 0.079155 .
## factor(stateid)SD -8.049e+00  2.498e+00 -3.222 0.001310 **
## factor(stateid)TN -3.120e+00  2.734e+00 -1.141 0.254056
## factor(stateid)TX  4.577e+00  2.800e+00  1.635 0.102368
## factor(stateid)UT -7.360e+00  2.549e+00 -2.888 0.003958 **
## factor(stateid)VA -4.008e+00  2.884e+00 -1.390 0.164868
## factor(stateid)VT -6.490e+00  2.712e+00 -2.393 0.016860 *
## factor(stateid)WA -6.748e+00  2.703e+00 -2.497 0.012680 *
## factor(stateid)WI -5.530e+00  2.680e+00 -2.063 0.039320 *
## factor(stateid)WV -3.118e+00  2.684e+00 -1.162 0.245618
## factor(stateid)WY -5.672e+00  2.624e+00 -2.162 0.030840 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.089 on 1091 degrees of freedom
## Multiple R-squared:  0.9815, Adjusted R-squared:  0.9805
## F-statistic: 980.2 on 59 and 1091 DF, p-value: < 2.2e-16
```

Sabit etkili bir model, bireyler (grup ve varlık) arasında aynı eğimleri ve sabit varyansı varsayarak veya aynı panel boyunca heterojenliğin sabit olduğunu varsayarak, kesişimlerdeki bireysel farklılıkları inceler.

Her bir devlet tanımlayıcısı tahmin için kukla değişken olarak kabul edilir.

Genel model, p-değeri 0.05'ten küçük olduğu için iyidir.

Değişkenler ve çoğu eyaletin bireysel etkisi önemlidir.

Düzeltilmiş  $R^2$ , havuzlanmış EKK'dan daha doğru ve iyi olan 0.9805'dir.

# Within Model Kullanılarak Sabit Etkiler

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = mur ~ vio + rob + incarc_rate + pop + pb1064 +
##      pw1064 + pm1029 + avginc + density, data = weapons, model = "within")
##
## Balanced Panel: n = 50, T = 23, N = 1150
##
## Residuals:
##      Min.      1st Qu.      Median      3rd Qu.      Max.
## -3.799410 -0.628962 -0.053542  0.605973  6.596966
##
## Coefficients:
##              Estimate Std. Error t-value Pr(>|t|)
## vio              0.00267833  0.00066951  4.0005 6.748e-05 ***
## rob              0.01636169  0.00159576 10.2532 < 2.2e-16 ***
## incarc_rate     -0.00696553  0.00073435 -9.4853 < 2.2e-16 ***
## pop             -0.41996141  0.06344867 -6.6189 5.663e-11 ***
## pb1064           0.63907169  0.13240384  4.8267 1.586e-06 ***
## pw1064           0.02826794  0.03616077  0.7817  0.4345
## pm1029           0.25521386  0.04628632  5.5138 4.382e-08 ***
## avginc           0.25187896  0.04617407  5.4550 6.058e-08 ***
## density          -8.67256486  4.43927460 -1.9536  0.0510 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    2500.7
## Residual Sum of Squares: 1294
## R-Squared:              0.48254
## Adj. R-Squared:         0.45504
## F-statistic: 113.044 on 9 and 1091 DF, p-value: < 2.22e-16
```

“Within” tahmin modeli, hataların karelerinin (SSE) toplamını düzeltir ancak hiçbir kukla değişken kullanılmadığından within sabit modelin hatalar için daha büyük serbestlik dereceleri vardır. Buna göre hata kareler ortalaması (MSE), tahminlerin standart hatalarını (SEE) veya hata kareleri ortalamasının karekökünü (RMSE) ve parametre tahminlerinin yanlış (daha küçük) standart hatalarını rapor eder.

Analiz sonuçlarına göre bu dönemde 10 ile 64 yaş arasındaki beyaz olan eyalet nüfusunun yüzdesi değişkeni hariç diğer değişkenlerin eyaletlerde cinayet oranı değeriyle anlamlı bir ilişkiye sahip olduğu gözlenmektedir.

10 ile 64 yaş arasındaki beyaz olan eyalet nüfusunun yüzdesi değişkeni hariç diğer değişkenlerin bireysel etkisi, p değeri 0.05’ten küçük olduğu için önemlidir.

R<sup>2</sup> ve Düzeltilmiş R<sup>2</sup>’nin değeri, havuzlanmış EKK ve LSDV’den daha küçük çıkmıştır.

# Bireysel Etkiler İçin F Testi

Bireysel etki için sabit etkili ve havuzlanmış EKK modelinin karşılaştırılması için F testi yapılır.

- $H_0$ : Havuzlanmış EKK modeli tutarlıdır.
- $H_a$ : Sabit etkili model tutarlıdır.

```
## F test for individual effects
##
## data:  mur ~ vio + rob + incarc_rate + pop + pb1064 + pw1064 + pm1029 + ...
## F = 53.44, df1 = 49, df2 = 1091, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

F testinin sonuçlarına göre,  $H_0$  hipotezi reddedilir. Böylelikle, sabit etkili modelin tutarlı olduğu gözlenmiştir.

```
## Oneway (individual) effect Random Effect Model
## (Swamy-Arora's transformation)
##
## Call:
## plm(formula = mur ~ vio + rob + incarc_rate + pop + pb1064 +
##      pw1064 + pm1029 + avginc + density, data = weapons, model = "random",
##      index = c("stateid", "year"))
##
## Balanced Panel: n = 50, T = 23, N = 1150
##
## Effects:
##
##              var std.dev share
## idiosyncratic 1.186   1.089 0.402
## individual    1.764   1.328 0.598
## theta: 0.8315
##
## Residuals:
##      Min.   1st Qu.   Median   3rd Qu.    Max.
## -4.39480 -0.69519 -0.11540  0.62055  6.60758
##
## Coefficients:
##              Estimate Std. Error z-value Pr(>|z|)
## (Intercept) -5.20481105  2.53782118 -2.0509 0.0402769 *
## vio          0.00368387  0.00066669  5.5256 3.283e-08 ***
## rob          0.01567103  0.00157398  9.9563 < 2.2e-16 ***
## incarc_rate -0.00601682  0.00072605 -8.2870 < 2.2e-16 ***
## pop         -0.12298410  0.03553989 -3.4605 0.0005393 ***
## pb1064       0.37525416  0.08476983  4.4267 9.567e-06 ***
## pw1064       0.03146574  0.03463766  0.9084 0.3636534
## pm1029       0.28135681  0.04703987  5.9812 2.214e-09 ***
## avginc       0.17981775  0.04177922  4.3040 1.677e-05 ***
## density     -5.63835117  0.93914500 -6.0037 1.929e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    2889.7
## Residual Sum of Squares: 1538
## R-Squared:              0.46776
## Adj. R-Squared:         0.46356
## Chisq: 1001.9 on 9 DF, p-value: < 2.22e-16
```

# Rastgele Etkili Model

Rastgele etkili modelde, heterojenlik her bir panelde farklılık gösterir. Rastgele etkili model, tahmin edilecek parametre sayısını azaltır ancak bireye özgü rastgele etki bağımsız değişkenlerle ilişkilendirildiğinde tutarsız tahminler üretecektir. Bu nedenle rastgele etkili model bireysel etkinin herhangi bir bağımsız değişkenlerle ilişkili olmadığını varsayar ve gruplara veya zamanlara özgü hata varyansları olarak tahmin edilir.

Analiz sonuçlarına göre bu dönemde 10 ile 64 yaş arasındaki beyaz olan eyalet nüfusunun yüzdesi değişkeni hariç diğer değişkenlerin eyaletlerde cinayet oranı değeriyle anlamlı bir ilişkiye sahip olduğu gözlenmektedir. Eyaletlerdeki bir önceki yıl eyalette hapsedilme oranı ve mil kare başına nüfus değişkeninin cinayet oranı üzerinde negatif etkisi olduğu gözlemlenmiştir.

10 ile 64 yaş arasındaki beyaz olan eyalet nüfusunun yüzdesi değişkeni hariç diğer değişkenlerin bireysel etkisi, p değeri 0.05'ten küçük olduğu için önemlidir.

$R^2$  ve Düzeltilmiş  $R^2$ 'nin değeri sırasıyla 0,46776 ve 0,46356'dır.



# Hausman Testi

Hausman testi, iki farklı modelin katsayı vektörlerinin farkına dayanmaktadır. Sabit veya Rastgele Etkili Model seçimine karar vermek için Hausman testi kullanılabilir.

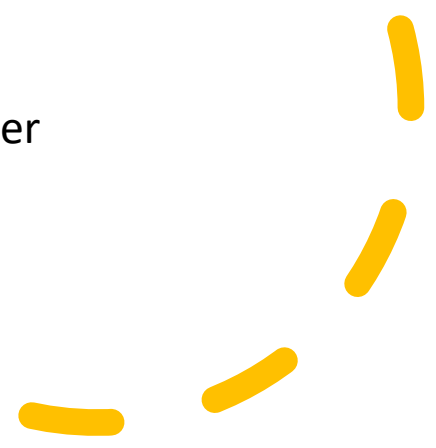
Sabit etkiler ve rassal etkiler tahminleyicileriyle tahminlenen modelden hangisini dikkate alacağımızı belirlemek amacıyla Hausman testi uygulanmıştır. Testin sonuçları aşağıda gösterilmektedir.

$H_0$ : Rastgele etkili model tutarlıdır.

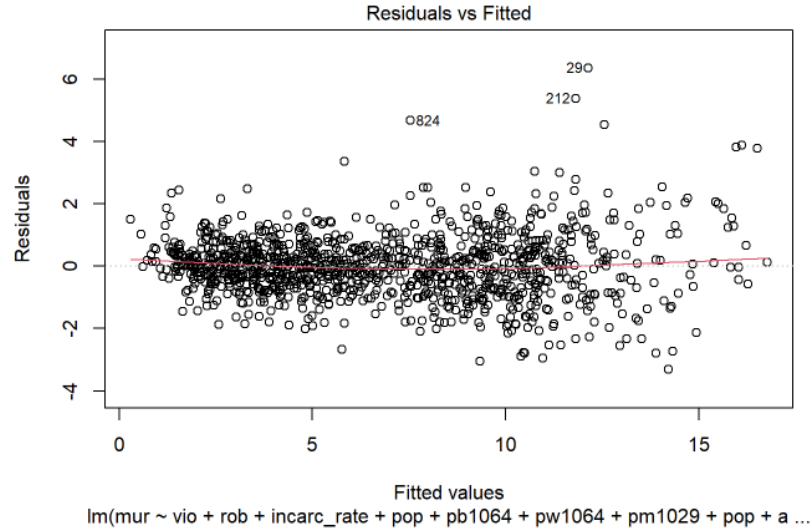
$H_a$ : Sabit etkili model tutarlıdır.

```
## Hausman Test
##
## data: mur ~ vio + rob + incarc_rate + pop + pb1064 + pw1064 + pm1029 + ...
## chisq = 634.92, df = 9, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
```

Görüldüğü üzere p değeri 0.05 den küçük olup istatistiksel olarak anlamlıdır. Bu nedenle sabit etkiler modelini dikkate almamız doğru olacaktır.



# Değişen Varyanslılık



Varyans sabit olmadığında doğrusal bir ilişki için değişen varyanslılık oluşur ve tahmin edici arttıkça artar. Bu gibi durumlarda, çıktıdaki standart hataya güvenilemez ancak yine de katsayılar tarafsız olacaktır.

Kesinlikle değişen varyans yoksa, X eksen aralığı boyunca tamamen rastgele, eşit bir nokta dağılımı ve düz bir kırmızı çizgi görülmelidir. Grafikler yanıltıcı olabilir bu nedenle Breusch-Pagan Test uygulanır.

```
##  
## Breusch-Pagan test  
##  
## data: mur ~ vio + rob + incarc_rate + pop + pb1064 + pw1064 + pm1029 + avginc + density + factor(stateid) + factor(y  
ear)  
## BP = 669.7, df = 80, p-value < 2.2e-16
```

$H_0$ : Değişen varyans problemi yoktur.

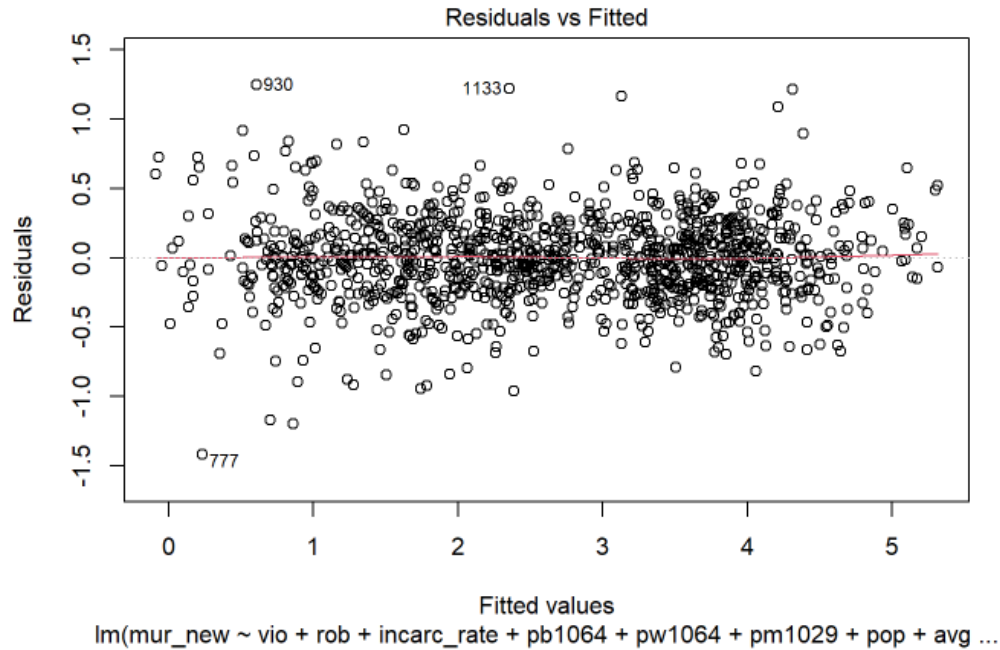
$H_a$ : Değişen varyans problemi vardır.

Breusch-Pagan (BP), lineer regresyon modelinin artıklarına bir lineer regresyon modeli uydurur ve varyansın çok fazla ek açıklayıcı değişkenler tarafından açıklanması durumunda reddeder.

BP testinin sonucuna göre p değeri 0.05 den küçük olup, tahmin değeri arttıkça kalıntıdaki varyansın değiştiğini gösterir, bu nedenle  $H_0$  hipotezi reddedilir. Değişkenler değişen varyansa sahiptir.

# Box-Cox Dönüşümü

Veriyi normal dağılıma yaklaştırmak için Box-Cox dönüşümü yapılır.



```
## studentized Breusch-Pagan test
##
## data: lmMod_bc
## BP = 212.39, df = 80, p-value = 6.073e-14
```

$H_0$ : Değişen varyans problemi yoktur.

$H_a$ : Değişen varyans problemi vardır.

BP testinin sonucuna göre p değeri 0.05' den küçüktür.  
Bu nedenle  $H_0$  hipotezi reddedilir. Değişkenler değişen varyansa sahiptir.

Görüldüğü üzere Box-Cox dönüşümünden sonra Residuals vs Fitted grafiği çok daha düz bir çizgiye sahiptir.

# Sonuç

---

- 1977-1999 yılları arasını kapsayan panel veride EKK yöntemi ile tahminleme yapılmıştır. Bu dönemde eyaletlerde cinayet oranındaki varyasyonu açıklaması bakımından eyalet nüfusu değişkeni hariç diğer değişkenlerin istatistiksel olarak anlamlı bir etkisi olduğu gözlemlenmiştir.
- Sabit etkili model kurulmuştur ve bu dönemde 10 ile 64 yaş arasındaki beyaz olan eyalet nüfusunun yüzdesi değişkeni hariç diğer değişkenlerin eyaletlerde cinayet oranı değeriyle anlamlı bir ilişkiye sahip olduğu gözlenmiştir.
- Rastgele etkili model kurulmuştur ve bu dönemde 10 ile 64 yaş arasındaki beyaz olan eyalet nüfusunun yüzdesi değişkeni hariç diğer değişkenlerin eyaletlerde cinayet oranı değeriyle anlamlı bir ilişkiye sahip olduğu gözlenmektedir.
- Hausman testi yapılarak sabit etkili modelin dikkate alınacağına karar verilmiştir.
- **Bu durumda, Amerika Birleşik Devletleri Cinayet Oranı değişkeni soygun oranı, şiddet içeren suç oranı, hapsedilme oranı, nüfus, 10 ile 29 yaşları arasındaki erkek eyalet nüfusunun yüzdesi, siyahi olan ve 10 ile 64 yaşları arasındaki eyalet nüfusunun yüzdesi, ortalama devlet geliri, mil kare başına nüfus değişkenlerinin anlamlı bir ilişkiye sahip olduğunu söyleyebiliriz.**



# Kaynakça

- <https://www.princeton.edu/~otorres/Panel101R.pdf>
- <https://cran.r-project.org/web/packages/sandwich/vignettes/sandwich.pdf>
- <https://www.r-bloggers.com/2016/01/how-to-detect-heteroscedasticity-and-rectify-it/>
- <https://uclspp.github.io/PUBL0055/seminar7.html>
- [https://uclspp.github.io/PUBL0055/lecture\\_slides/lecture7.pdf](https://uclspp.github.io/PUBL0055/lecture_slides/lecture7.pdf)
- <https://www.econometrics-with-r.org/10.1-panel-data.html>
- <https://www.istmer.com/panel-veri-analizi-nedir/>
- Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, Cilt: 32 2018 Sayı: 4 (PANEL VERİ ANALİZİ VE KÜMELEME YÖNTEMİ İLE TÜRKİYE'DE KONUT TALEBİNİN İNCELENMESİ) -Cahit ÇELİK -Gülşen KIRAL <https://dergipark.org.tr/tr/download/article-file/548380>
- İstanbul Üniversitesi AUZEF - "Panel Veri Analizi" -PROF. DR. FERDA YERDELEN TATOĞLU [http://auzefkitap.istanbul.edu.tr/kitap/ekonometri\\_ue/panelverianalizi.pdf](http://auzefkitap.istanbul.edu.tr/kitap/ekonometri_ue/panelverianalizi.pdf)