Example of Poisson Regression in Section2.3:

```
bike      Number of Deaths by Bicycle Accident (自転車事故死者数, 2012)
lowland   Lowland Area (低地面積, 平方キロ, 2012)
residen   Residential Land Area (居住用宅地面積, 平方キロ, 2012)
pop       Population (2010)
```

	pref	bike	lowland	dwellings	pop
北海道	1	11	9794	543	5504
青森	2	6	1237	193	1374
岩手	3	7	1261	216	1326
宮城	4	4	1757	259	2352
秋田	5	2	2453	170	1085
山形	6	5	1393	163	1167
福島	7	5	1437	255	2021
茨城	8	20	1647	454	2887
栃木	9	17	752	289	1990
群馬	10	17	585	272	2005
埼玉	11	42	1414	487	6373
千葉	12	30	1452	489	5560
東京	13	34	274	421	15576
神奈川	14	17	575	418	8254
新潟	19	5	2775	274	2375
富山	20	4	987	145	1091
石川	15	5	656	116	1172
福井	16	2	932	93	807
山梨	17	4	343	115	855
長野	21	7	751	307	2149

岐阜	22	12	1174	226	1998
静岡	23	22	1155	338	3760
愛知	24	44	1148	521	7521
三重	18	8	1031	207	1820
滋賀	25	6	935	132	1363
京都	26	15	820	149	2668
大阪	27	47	610	318	9281
兵庫	28	23	1604	346	5348
奈良	29	4	273	110	1260
和歌山	30	7	316	93	983
鳥取	31	4	411	70	589
島根	32	3	495	94	718
岡山	33	14	1141	216	1943
広島	34	12	559	232	2869
山口	35	2	461	173	1444
徳島	36	7	551	88	783
香川	37	17	474	117	998
愛媛	38	9	557	146	1433
高知	39	6	327	70	763
福岡	40	18	1224	400	5078
佐賀	41	6	645	103	852
長崎	42	1	339	141	1423
熊本	43	14	958	225	1810
大分	44	6	595	140	1197
宮崎	45	6	764	163	1136
鹿児島	46	5	771	258	1704
沖縄	47	1	151	98	1392

```
----

. poisson bike lowland residen pop

Iteration 0:   log likelihood = -156.83031
Iteration 1:   log likelihood = -153.97721
Iteration 2:   log likelihood = -153.97403
Iteration 3:   log likelihood = -153.97403

Poisson regression                                Number of obs   =         47
                                                  LR chi2(3)      =     286.85
                                                  Prob > chi2     =     0.0000
Log likelihood = -153.97403                       Pseudo R2       =     0.4823

------------------------------------------------------------------------------
        bike |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     lowland |  -.0001559   .0000368    -4.23   0.000    -.0002281   -.0000837
     residen |   .0042478    .000447     9.50   0.000     .0033716    .0051239
         pop |   .0000519   .0000146     3.56   0.000     .0000234    .0000804
       _cons |   1.309844   .1051302    12.46   0.000     1.103793    1.515896
------------------------------------------------------------------------------

. gen llland=log(lowland)

. gen lresiden=log(residen)

. gen lpop=log(pop)
```

```
. poisson bike llland lresiden lpop

Iteration 0:   log likelihood = -156.15686
Iteration 1:   log likelihood = -155.62550
Iteration 2:   log likelihood = -155.62489
Iteration 3:   log likelihood = -155.62489

Poisson regression                                Number of obs   =         47
                                                  LR chi2(3)      =     283.54
                                                  Prob > chi2     =     0.0000
Log likelihood = -155.62489                       Pseudo R2       =     0.4767

------------------------------------------------------------------------------
        bike |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      llland |  -.1028579   .0800629    -1.28   0.199    -.2597784    .0540625
    lresiden |   .4817018   .2171779     2.22   0.027      .056041    .9073626
        lpop |   .5715923   .1220733     4.68   0.000      .332333    .8108517
       _cons |   -3.93974    .559487    -7.04   0.000    -5.036315   -2.843166
------------------------------------------------------------------------------
```

3 Panel Data

3.1 GLS — Review

Regression model I:

$$y = X\beta + u, \qquad u \sim N(0, \sigma^2 I_n),$$

where y, X, β, u, 0 and I_n are $n \times 1$, $n \times k$, $k \times 1$, $n \times 1$, $n \times 1$, and $n \times n$, respectively.

We solve the following minimization problem:

$$\min_{\beta} (y - X\beta)'(y - X\beta).$$

Let $\hat{\beta}$ be a solution of the above minimization problem.

OLS estimator of β is given by:

$$\hat{\beta} = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'u.$$

$$\mathrm{E}(\hat{\beta}) = \beta, \qquad \mathrm{V}(\hat{\beta}) = \sigma^2(X'X)^{-1}.$$

Regression model II:

$$y = X\beta + u, \qquad u \sim N(0, \sigma^2\Omega),$$

where Ω is $n \times n$.

We solve the following minimization problem:

$$\min_{\beta} (y - X\beta)'\Omega^{-1}(y - X\beta).$$

Let b be a solution of the above minimization problem.

GLS estimator of β is given by:

$$b = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y = \beta + (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}u.$$

$$\mathrm{E}(b) = \beta, \qquad \mathrm{V}(b) = \sigma^2(X'\Omega^{-1}X)^{-1}.$$

- We apply OLS to the following regression model:

$$y = X\beta + u, \qquad u \sim N(0, \sigma^2\Omega).$$

OLS estimator of β is given by:

$$\hat{\beta} = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'u.$$

$$\mathrm{E}(\hat{\beta}) = \beta, \qquad \mathrm{V}(\hat{\beta}) = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}.$$

$\hat{\beta}$ is an unbiased estimator.

The difference between two variances is:

$$
\begin{aligned}
&\mathrm{V}(\hat{\beta}) - \mathrm{V}(b) \\
&= \sigma^2 (X'X)^{-1} X'\Omega X (X'X)^{-1} - \sigma^2 (X'\Omega^{-1}X)^{-1} \\
&= \sigma^2 \Big((X'X)^{-1}X' - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1} \Big) \Omega \Big((X'X)^{-1}X' - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1} \Big)' \\
&= \sigma^2 A\Omega A'
\end{aligned}
$$

Ω is the variance-covariance matrix of u, which is a positive definite matrix.

Therefore, except for $\Omega = I_n$, $A\Omega A'$ is also a positive definite matrix.

This implies that $\mathrm{V}(\hat{\beta}_i) - \mathrm{V}(b_i) > 0$ for the ith element of β.

Accordingly, b is more efficient than $\hat{\beta}$.

3.2 Panel Model Basic

Model:

$$y_{it} = X_{it}\beta + v_i + u_{it}, \qquad i = 1, 2, \cdots, n, \quad t = 1, 2, \cdots, T$$

where i indicates individual and t denotes time.

There are n observations for each t.

u_{it} indicates the error term, assuming that $\mathrm{E}(u_{it}) = 0$, $\mathrm{V}(u_{it}) = \sigma_u^2$ and $\mathrm{Cov}(u_{it}, u_{js}) = 0$ for $i \neq j$ and $t \neq s$.

v_i denotes the individual effect, which is fixed or random.

3.2.1 Fixed Effect Model (固定効果モデル)

In the case where v_i is fixed, the case of $v_i = z_i\alpha$ is included.

$$y_{it} = X_{it}\beta + v_i + u_{it}, \qquad i = 1, 2, \cdots, n, \quad t = 1, 2, \cdots, T,$$

$$\overline{y}_i = \overline{X}_i\beta + v_i + \overline{u}_i, \qquad i = 1, 2, \cdots, n,$$

where $\overline{y}_i = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T} y_{it}$, $\overline{X}_i = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T} X_{it}$, and $\overline{u}_i = \dfrac{1}{T}\displaystyle\sum_{t=1}^{T} u_{it}$.

$$(y_{it} - \overline{y}_i) = (X_{it} - \overline{X}_i)\beta + (u_{it} - \overline{u}_i), \qquad i = 1, 2, \cdots, n, \quad t = 1, 2, \cdots, T,$$

Taking an example of y, the left-hand side of the above equation is rewritten as:

$$y_{it} - \overline{y}_i = y_{it} - \frac{1}{T}1'_T y_i,$$

where $1_T = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$, which is a $T \times 1$ vector, and $y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix}$.

$$\begin{pmatrix} y_{i1} - \overline{y}_i \\ y_{i2} - \overline{y}_i \\ \vdots \\ y_{iT} - \overline{y}_i \end{pmatrix} = I_T y_i - 1_T \overline{y}_i = I_T y_i - \frac{1}{T} 1_T 1_T' y_i = (I_T - \frac{1}{T} 1_T 1_T') y_i$$

Thus,

$$\begin{pmatrix} y_{i1} - \overline{y}_i \\ y_{i2} - \overline{y}_i \\ \vdots \\ y_{iT} - \overline{y}_i \end{pmatrix} = \begin{pmatrix} X_{i1} - \overline{X}_i \\ X_{i2} - \overline{X}_i \\ \vdots \\ X_{iT} - \overline{X}_i \end{pmatrix} \beta + \begin{pmatrix} u_{i1} - \overline{u}_i \\ u_{i2} - \overline{u}_i \\ \vdots \\ u_{iT} - \overline{u}_i \end{pmatrix}, \qquad i = 1, 2, \cdots, n,$$

which is re-written as:

$$(I_T - \frac{1}{T}1_T1_T')y_i = (I_T - \frac{1}{T}1_T1_T')X_i\beta + (I_T - \frac{1}{T}1_T1_T')u_i, \qquad i = 1, 2, \cdots, n,$$

i.e.,

$$D_Ty_i = D_TX_i\beta + D_Tu_i, \qquad i = 1, 2, \cdots, n,$$

where $D_T = (I_T - \frac{1}{T}1_T1_T')$, which is a $T \times T$ matrix.

Note that $D_TD_T' = D_T$, i.e., D_T is a symmetric and idempotent matrix.

Using the matrix form for $i = 1, 2, \cdots, n$, we have:

$$\begin{pmatrix} D_Ty_1 \\ D_Ty_2 \\ \vdots \\ D_Ty_n \end{pmatrix} = \begin{pmatrix} D_TX_1 \\ D_TX_2 \\ \vdots \\ D_TX_n \end{pmatrix}\beta + \begin{pmatrix} D_Tu_1 \\ D_Tu_2 \\ \vdots \\ D_Tu_n \end{pmatrix},$$

i.e.,

$$\begin{pmatrix} D_T & 0 & \cdots & 0 \\ 0 & D_T & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_T \end{pmatrix} y = \begin{pmatrix} D_T & 0 & \cdots & 0 \\ 0 & D_T & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_T \end{pmatrix} X\beta + \begin{pmatrix} D_T & 0 & \cdots & 0 \\ 0 & D_T & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & D_T \end{pmatrix} u,$$

where $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$, and $u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$, which are $Tn \times 1$, $Tn \times k$ and $Tn \times 1$ matrices, respectively

Using the Kronecker product, we obtain the following expression:

$$(I_n \otimes D_T)y = (I_n \otimes D_T)X\beta + (I_n \otimes D_T)u,$$

where $(I_n \otimes D_T)$, y, X, and u are $nT \times nT$, $nT \times 1$, $nT \times k$, and $nT \times 1$, respectively.