Zero-Inflated Poisson Count Data Model: In the case of too many zeros, we have to modify the estimation procedure.

Suppose that the probability of $y_i = j$ is decomposed of two regimes.

$\longrightarrow$ We have the case of $y_i = j$ and Regime 1, and that of $y_i = j$ and Regime 2.

Consider $P(y_i = 0)$ and $P(y_i = j)$ separately as follows:

$$P(y_i = 0) = P(y_i = 0|\text{Regime 1})P(\text{Regime 1}) + P(y_i = 0|\text{Regime 2})P(\text{Regime 2})$$

$$P(y_i = j) = P(y_i = j|\text{Regime 1})P(\text{Regime 1}) + P(y_i = j|\text{Regime 2})P(\text{Regime 2}),$$

for $j = 1, 2, \cdots$.

Assume:

- $P(y_i = 0|\text{Regime 1}) = 1$ and $P(y_i = j|\text{Regime 1}) = 0$ for $j = 1, 2, \cdots$,
- $P(\text{Regime 1}) = F_i$ and $P(\text{Regime 2}) = 1 - F_i$,
- $P(y_i = j|\text{Regime 2}) = \dfrac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$ for $j = 0, 1, 2, \cdots$,

where $F_i = F(Z_i\alpha)$ and $\lambda_i = \exp(X_i\beta)$. $\Longrightarrow$ Z_i and X_i are exogenous variables.

Under the first assumption, we have the following equations:

$$P(y_i = 0) = P(\text{Regime 1}) + P(y_i = 0|\text{Regime 2})P(\text{Regime 2})$$

$$P(y_i = j) = P(y_i = j|\text{Regime 2})P(\text{Regime 2}),$$

for $j = 1, 2, \cdots$.

Combining the above two equations, we obtain the following:

$$P(y_i = j) = P(\text{Regime 1})I_i + P(y_i = j|\text{Regime 2})P(\text{Regime 2}),$$

for $j = 0, 1, 2, \cdots$,

where the indicator function I_i is given by $I_i = 1$ for $y_i = 0$ and $I_i = 0$ for $y_i \neq 0$.

F_i denotes a cumulative distribution function of $Z_i\alpha$. $\Longrightarrow$ We have to assume F_i.

Including the other two assumptions, we obtain the distribution of y_i as follows:

$$P(y_i = j) = F_i I_i + \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}(1 - F_i), \qquad j = 0, 1, 2, \cdots$$

where $F_i \equiv F(Z_i\alpha)$, $\lambda_i = \exp(X_i\beta)$, and the indicator function I_i is given by $I_i = 1$ for $y_i = 0$ and $I_i = 0$ for $y_i \neq 0$.

Therefore, the log-likelihood function is:

$$\log L(\alpha,\beta) = \sum_{i=1}^{n} \log P(y_i = j) = \sum_{i=1}^{n} \log\left(F_i I_i + \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}(1 - F_i)\right),$$

where $F_i \equiv F(Z_i\alpha)$ and $\lambda_i = \exp(X_i\beta)$.

$\log L(\alpha,\beta)$ is maximized with respect to α and β.

$\Longrightarrow$ The Newton-Raphson method or the method of scoring is utilized for maximization.

2.4 Applications

Empirical Application of Example 1 in Section 2.1: Excess Demand Function.

Demand is observed as both amount and quantity, while supply is not.

Therefore, excess demand is not observed,

Data are taken from household expenditure survey as follows:

```
y    Income (実収入, Japanese Yen)
s    Sake (清酒, Japanese Yen)
b    Beer (ビール, Japanese Yen)
sml  Sake (清酒, 1ml)
bl   Beer (ビール, 1l)
CPI  Consumer Price Index (消費者物価指数・総合, Base Year 2010)
```

year	y	s	b	sml	bl	CPI
2000.01	458911	716	1350	828	2.67	102.8
2000.02	486601	643	1527	728	3.01	102.5
2000.03	494395	661	1873	775	3.69	102.7
2000.04	505409	614	1967	749	3.93	102.9
2000.05	460116	567	2311	679	4.64	103.0
2000.06	772611	518	2225	596	4.40	102.8
2000.07	640258	459	3419	511	6.57	102.5
2000.08	506757	455	2976	530	5.91	102.8
2000.09	446405	477	2160	580	4.27	102.7
2000.10	488921	626	1805	750	3.59	102.7
2000.11	457054	680	1674	831	3.36	102.4
2000.12	1035616	1623	2546	1688	5.04	102.5
2001.01	453748	689	1363	806	2.71	102.5
2001.02	475556	554	1299	688	2.59	102.1
2001.03	481198	567	1467	708	2.99	101.9
2001.04	498080	532	1641	637	3.33	102.1
2001.05	447510	486	1825	608	3.63	102.2
2001.06	766471	446	2003	535	3.99	101.9
2001.07	614715	493	2656	568	5.25	101.6
2001.08	496482	436	2326	492	4.60	102.0
2001.09	447397	479	1546	617	3.06	101.8
2001.10	489834	568	1426	733	2.87	101.8
2001.11	461094	646	1222	818	2.42	101.3
2001.12	1000728	1609	2274	1710	4.54	101.2
2002.01	462389	637	1040	716	2.01	101.0
2002.02	477622	570	1040	778	2.15	100.5

2002.03	496351	552	1418	748	2.81	100.7
2002.04	485770	502	1427	689	2.93	101.0
2002.05	444612	497	1623	602	3.40	101.3
2002.06	745480	442	1900	537	3.74	101.2
2002.07	583862	499	2437	554	4.72	100.8
2002.08	488257	472	2358	508	4.72	101.1
2002.09	440319	437	1522	536	2.98	101.1
2002.10	475494	561	1378	757	2.85	100.9
2002.11	439186	730	1347	888	2.59	100.9
2002.12	939747	1589	2177	1936	4.21	100.9
2003.01	435989	549	1025	632	1.98	100.6
2003.02	455309	519	1089	670	2.19	100.3
2003.03	456873	531	1343	686	2.55	100.6
2003.04	475037	514	1369	576	2.69	100.9
2003.05	429669	518	1396	724	2.73	101.1
2003.06	730617	484	1609	597	3.17	100.8
2003.07	574574	492	2013	636	3.93	100.6
2003.08	474973	503	2146	641	4.33	100.8
2003.09	429301	395	1331	463	2.65	100.9
2003.10	467408	498	1312	560	2.64	100.9
2003.11	435079	560	1230	760	2.42	100.4
2003.12	932887	1484	2012	1621	3.97	100.5
2004.01	445133	530	1062	595	2.10	100.3
2004.02	474143	591	1086	705	2.20	100.3
2004.03	456288	455	1239	621	2.43	100.5
2004.04	488217	441	1273	539	2.47	100.5
2004.05	446758	438	1530	524	3.06	100.6
2004.06	723370	391	1729	447	3.37	100.8
2004.07	599045	414	2166	432	4.18	100.5

2004.08	476264	403	2032	474	4.05	100.6
2004.09	440187	387	1414	450	2.82	100.9
2004.10	467895	454	1269	551	2.54	101.4
2004.11	442885	482	1266	619	2.54	101.2
2004.12	920100	1262	1912	1272	3.83	100.7
2005.01	448635	542	999	678	1.94	100.5
2005.02	469673	497	917	630	1.84	100.2
2005.03	451360	485	1060	714	2.05	100.5
2005.04	495036	406	1226	505	2.47	100.6
2005.05	440388	386	1437	443	2.84	100.7
2005.06	720667	375	1472	430	2.96	100.3
2005.07	576129	451	2214	555	4.36	100.2
2005.08	463034	370	2001	440	4.04	100.3
2005.09	427753	323	1321	390	2.56	100.6
2005.10	463838	500	1246	624	2.39	100.6
2005.11	433036	519	1064	678	2.12	100.2
2005.12	905473	1173	2090	1152	4.06	100.3
2006.01	437787	466	921	501	1.82	100.4
2006.02	461368	433	884	580	1.71	100.1
2006.03	429948	416	1060	517	2.06	100.3
2006.04	472583	444	1269	536	2.43	100.5
2006.05	426680	426	1367	544	2.55	100.8
2006.06	684632	431	1360	529	2.60	100.8
2006.07	613269	358	1803	395	3.47	100.5
2006.08	475866	400	1843	448	3.50	101.2
2006.09	429017	341	1139	444	2.20	101.2
2006.10	467163	479	1183	696	2.35	101.0
2006.11	442147	533	1053	660	2.01	100.5
2006.12	968162	1144	1882	1200	3.76	100.6

2007.01	441039	505	941	695	1.82	100.4
2007.02	471681	428	899	580	1.69	99.9
2007.03	445076	434	1071	528	2.07	100.2
2007.04	472446	413	1291	506	2.60	100.5
2007.05	431013	346	1302	450	2.42	100.8
2007.06	735579	374	1532	490	2.93	100.6
2007.07	592452	414	1845	530	3.63	100.5
2007.08	467786	368	2121	511	4.10	101.0
2007.09	431793	329	1446	425	2.80	101.0
2007.10	469981	445	1108	542	2.15	101.3
2007.11	435640	541	1116	594	2.20	101.1
2007.12	950654	1085	1892	1209	3.56	101.3
2008.01	438998	509	1000	707	1.99	101.1
2008.02	476282	445	1008	558	1.98	100.9
2008.03	453482	400	1199	573	2.35	101.4
2008.04	469774	376	1234	492	2.44	101.3
2008.05	435076	329	1404	406	2.72	102.1
2008.06	737166	356	1410	395	2.72	102.6
2008.07	587732	298	1832	338	3.48	102.8
2008.08	488216	334	1767	413	3.36	103.1
2008.09	433502	293	1086	423	2.03	103.1
2008.10	481746	346	1066	434	2.04	103.0
2008.11	439394	439	1077	533	2.06	102.1
2008.12	969449	1076	1711	1231	3.24	101.7
2009.01	443337	479	962	636	1.85	101.1
2009.02	464665	417	849	705	1.64	100.8
2009.03	443429	444	1009	478	1.88	101.1
2009.04	473779	354	958	428	1.87	101.2
2009.05	436123	370	1180	495	2.34	101.0

2009.06	700239	343	1126	386	2.14	100.8
2009.07	573821	287	1478	327	2.78	100.5
2009.08	466393	300	1519	345	2.90	100.8
2009.09	422120	263	974	363	1.86	100.8
2009.10	459704	349	941	435	1.81	100.4
2009.11	428219	432	941	588	1.81	100.2
2009.12	906884	943	1546	1019	2.98	100.0
2010.01	434344	420	800	464	1.46	100.1
2010.02	464866	347	751	500	1.49	100.0
2010.03	439410	386	885	578	1.74	100.3
2010.04	474616	317	926	404	1.79	100.4
2010.05	421413	316	1040	455	1.99	100.3
2010.06	733886	316	1236	375	2.36	100.1
2010.07	562094	362	1600	382	3.05	99.5
2010.08	470717	314	1571	397	3.03	99.7
2010.09	425771	255	1028	361	1.93	99.9
2010.10	494398	337	1017	520	1.95	100.2
2010.11	431281	374	870	485	1.67	99.9
2010.12	895511	943	1456	912	2.80	99.6
2011.01	419728	418	693	495	1.33	99.5
2011.02	470071	345	650	552	1.23	99.5
2011.03	419862	366	703	472	1.34	99.8
2011.04	454433	371	814	485	1.53	99.9
2011.05	413506	345	888	432	1.67	99.9
2011.06	687212	317	1025	327	1.95	99.7
2011.07	572662	267	1407	367	2.68	99.7
2011.08	463760	277	1378	345	2.66	99.9
2011.09	422720	276	917	419	1.77	99.9
2011.10	479749	345	789	433	1.52	100.0

2011.11	424272	329	848	426	1.64	99.4
2011.12	893811	884	1398	907	2.73	99.4
2012.01	430477	432	711	619	1.43	99.6
2012.02	483625	394	721	495	1.39	99.8
2012.03	441015	397	787	592	1.50	100.3
2012.04	469381	381	833	466	1.56	100.4
2012.05	417723	309	845	411	1.67	100.1
2012.06	712592	337	1015	417	1.96	99.6
2012.07	557032	284	1242	375	2.36	99.3
2012.08	470470	288	1374	357	2.61	99.4
2012.09	422046	294	903	337	1.76	99.6
2012.10	482101	282	752	361	1.45	99.6
2012.11	432681	361	756	510	1.40	99.2
2012.12	902928	859	1347	863	2.56	99.3
2013.01	433858	377	743	467	1.43	99.3
2013.02	476256	325	656	410	1.26	99.2
2013.03	444379	384	815	467	1.52	99.4
2013.04	479854	323	680	359	1.21	99.7
2013.05	422724	322	853	402	1.61	99.8
2013.06	728678	433	973	541	1.86	99.8
2013.07	569174	281	1104	315	2.04	100.0
2013.08	471411	298	1200	324	2.32	100.3
2013.09	431931	258	848	311	1.59	100.6
2013.10	482684	282	805	296	1.47	100.7
2013.11	436293	377	725	447	1.32	100.8
2013.12	905822	835	1351	933	2.62	100.9
2014.01	438646	431	703	530	1.38	100.7
2014.02	479268	365	612	446	1.21	100.7
2014.03	438145	397	891	476	1.66	101.0

```
2014.04  463964  304  630  401 1.15 103.1
2014.05  421117  348  846  432 1.56 103.5
2014.06  710375  356  933  394 1.72 103.4
2014.07  555276  304 1182  361 2.18 103.4
2014.08  463810  325 1159  356 2.12 103.7
2014.09  421809  319  840  383 1.51 103.9
2014.10  488273  345  707  391 1.33 103.6
2014.11  431543  398  716  519 1.33 103.2
2014.12  924911  892 1324  901 2.47 103.3
2015.01  440226  400  622  494 1.14 103.1
2015.02  488519  356  600  469 1.12 102.9
2015.03  449243  353  712  416 1.28 103.3
2015.04  476880  353  739  413 1.35 103.7
2015.05  430325  331  909  377 1.64 104.0
2015.06  733589  350  928  231 1.66 103.8
2015.07  587156  331 1105  347 2.02 103.7
2015.08  475369  339 1165  462 2.18 103.9
2015.09  415467  346  797  354 1.47 103.9


. gen t=_n                              <---- make data

. tsset t                               <---- set t as time series data
        time variable:  t, 1 to 189
                delta:  1 unit

. gen ry=y/(cpi/100)                    <---- real income
```

```
. gen rsp=(s/sml)/(cpi/100)        <---- real sake price per 1ml

. gen rbp=(b/bl)/(cpi/100)         <---- real beer price per 1l

. gen ds=0                         <---- dfault data

. replace ds=1 if f.rsp>rsp        <---- set ds=1 when excess demand exists
(94 real changes made)

. probit ds ry rsp rbp, if t<188.5 <---- Estimate probit
                                         during the period from 1 to 188

Iteration 0:   log likelihood = -130.30103
Iteration 1:   log likelihood = -95.883766
Iteration 2:   log likelihood = -95.419505
Iteration 3:   log likelihood = -95.419207
Iteration 4:   log likelihood = -95.419207

Probit regression                                 Number of obs   =        188
                                                  LR chi2(3)      =      69.76
                                                  Prob > chi2     =     0.0000
Log likelihood = -95.419207                       Pseudo R2       =     0.2677

------------------------------------------------------------------------------
          ds |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ry |   8.04e-07   9.42e-07     0.85   0.393    -1.04e-06    2.65e-06
         rsp |   -13.8574   2.166711    -6.40   0.000    -18.10408   -9.610729
         rbp |   .0026681   .0067234     0.40   0.691    -.0105094    .0158457
```

```
       _cons |    9.44494   3.697318     2.55   0.011     2.198331    16.69155
------------------------------------------------------------------------------
Note: 1 failure and 0 successes completely determined.

. logit ds ry rsp rbp if t<188.5 <---- Estimate logit
                                       during the period from 1 to 188

Iteration 0:   log likelihood = -130.30103
Iteration 1:   log likelihood = -96.132508
Iteration 2:   log likelihood =  -95.65503
Iteration 3:   log likelihood = -95.653538
Iteration 4:   log likelihood = -95.653538

Logistic regression                               Number of obs   =        188
                                                  LR chi2(3)      =      69.29
                                                  Prob > chi2     =     0.0000
Log likelihood = -95.653538                       Pseudo R2       =     0.2659

------------------------------------------------------------------------------
          ds |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ry |   1.41e-06   1.56e-06     0.90   0.368    -1.65e-06    4.47e-06
         rsp |  -23.36485   3.941076    -5.93   0.000    -31.08922   -15.64048
         rbp |   .0046687   .0113128     0.41   0.680    -.0175041    .0268414
       _cons |   15.82621   6.301665     2.51   0.012     3.475172    28.17724
------------------------------------------------------------------------------
```

$$D_t - S_t = \beta_0 + \beta_1 ry_t + \beta_2 rsp_t + \beta_3 rbp_t$$

D_t is observed, but S_t is not observed. Therefore, $D_t - S_t$ is unobserved.

$$rsp_{t+1} > rsp_t \implies D_t - S_t > 0 \implies ds_t = 1.$$

$$rsp_{t+1} \le rsp_t \implies D_t - S_t \le 0 \implies ds_t = 0.$$

Example of Truncated Regression Model in Section 2.2: Demand Function of Watermelon

```
Workers' Households (二人以上の世帯のうち勤労者世帯) (2000 -)

y    Income  【Japanese Yen】
a    Apple   【Japanese Yen】
g    Grape   【Japanese Yen】
w    Watermelon  【Japanese Yen】
ag   Apple   【1g】
gg   Grape   【1g】
wg   Watermelon  【1g】
CPI  Consumer Price Index (Base year 2010)
```

year	y	a	g	w	ag	gg	wg	CPI
2000.01	458911	371	6	3	1093	9	3	102.8
2000.02	486601	416	4	4	1285	5	4	102.5
2000.03	494395	388	8	8	1145	12	10	102.7
2000.04	505409	350	19	46	899	28	85	102.9
2000.05	460116	258	46	243	598	45	638	103.0
2000.06	772611	191	153	352	446	169	1163	102.8
2000.07	640258	139	317	571	306	293	2152	102.5
2000.08	506757	144	1032	397	282	1073	1558	102.8
2000.09	446405	354	826	30	884	1002	109	102.7
2000.10	488921	501	292	3	1460	360	8	102.7
2000.11	457054	739	37	1	2024	43	2	102.4

2000.12	1035616	938	16	5	2230	27	11	102.5
2001.01	453748	329	11	1	905	16	2	102.5
2001.02	475556	350	5	1	920	6	3	102.1
2001.03	481198	321	7	3	835	11	5	101.9
2001.04	498080	287	17	52	713	26	92	102.1
2001.05	447510	255	43	236	582	43	602	102.2
2001.06	766471	169	138	355	352	120	1167	101.9
2001.07	614715	108	301	616	203	278	2403	101.6
2001.08	496482	129	827	400	265	916	1577	102.0
2001.09	447397	449	661	26	1087	823	90	101.8
2001.10	489834	598	241	1	1581	308	2	101.8
2001.11	461094	673	27	1	2026	34	2	101.3
2001.12	1000728	961	16	3	2622	16	6	101.2
2002.01	462389	331	4	2	997	4	3	101.0
2002.02	477622	343	2	1	1327	2	1	100.5
2002.03	496351	326	8	8	1114	10	22	100.7
2002.04	485770	273	14	50	826	21	90	101.0
2002.05	444612	243	57	208	726	55	517	101.3
2002.06	745480	194	170	353	524	157	1225	101.2
2002.07	583862	126	324	499	313	341	2075	100.8
2002.08	488257	151	722	335	312	813	1406	101.1
2002.09	440319	376	730	24	939	853	88	101.1
2002.10	475494	506	366	1	1504	462	3	100.9
2002.11	439186	733	36	3	2056	52	3	100.9
2002.12	939747	847	24	2	2599	38	2	100.9
2003.01	435989	303	7	1	900	12	0	100.6
2003.02	455309	305	3	2	1148	5	1	100.3
2003.03	456873	326	11	2	1094	22	8	100.6
2003.04	475037	273	18	36	815	28	63	100.9

2003.05	429669	221	40	171	583	58	422	101.1
2003.06	730617	157	177	294	368	150	967	100.8
2003.07	574574	153	244	379	326	242	1412	100.6
2003.08	474973	128	683	293	264	873	1110	100.8
2003.09	429301	333	636	33	938	738	88	100.9
2003.10	467408	506	258	5	1193	346	5	100.9
2003.11	435079	618	39	1	2105	46	0	100.4
2003.12	932887	757	12	3	1856	13	2	100.5
2004.01	445133	327	5	1	995	6	0	100.3
2004.02	474143	348	3	2	1044	4	3	100.3
2004.03	456288	287	7	5	829	9	6	100.5
2004.04	488217	221	13	52	640	26	114	100.5
2004.05	446758	192	52	168	487	61	542	100.6
2004.06	723370	141	133	289	362	123	725	100.8
2004.07	599045	94	313	462	223	307	1689	100.5
2004.08	476264	115	675	276	260	761	892	100.6
2004.09	440187	328	583	25	859	814	82	100.9
2004.10	467895	482	156	1	1192	204	4	101.4
2004.11	442885	563	48	2	1613	58	7	101.2
2004.12	920100	673	15	3	1686	24	0	100.7
2005.01	448635	310	6	4	785	9	3	100.5
2005.02	469673	340	4	6	911	4	17	100.2
2005.03	451360	360	11	7	933	13	9	100.5
2005.04	495036	294	18	23	787	30	37	100.6
2005.05	440388	226	47	149	485	50	416	100.7
2005.06	720667	152	126	337	335	111	1088	100.3
2005.07	576129	105	217	402	216	226	1546	100.2
2005.08	463034	104	582	328	234	652	1225	100.3
2005.09	427753	277	644	30	771	838	93	100.6

2005.10	463838	404	363	1	1147	544	4	100.6
2005.11	433036	540	45	1	1594	67	1	100.2
2005.12	905473	631	13	2	1519	20	2	100.3
2006.01	437787	294	7	1	994	10	1	100.4
2006.02	461368	310	4	0	950	6	0	100.1
2006.03	429948	302	7	7	920	12	0	100.3
2006.04	472583	256	17	25	728	26	40	100.5
2006.05	426680	202	32	141	515	44	332	100.8
2006.06	684632	148	114	240	338	97	720	100.8
2006.07	613269	105	209	361	228	205	1413	100.5
2006.08	475866	82	595	324	163	634	1034	101.2
2006.09	429017	263	628	32	647	716	108	101.2
2006.10	467163	455	263	4	1144	359	4	101.0
2006.11	442147	605	23	0	1556	22	1	100.5
2006.12	968162	719	18	1	1949	13	0	100.6
2007.01	441039	309	5	1	858	4	0	100.4
2007.02	471681	319	3	8	950	5	0	99.9
2007.03	445076	346	6	2	1012	9	0	100.2
2007.04	472446	304	15	35	770	23	75	100.5
2007.05	431013	233	35	159	539	37	355	100.8
2007.06	735579	177	122	320	369	110	926	100.6
2007.07	592452	110	201	360	258	212	1322	100.5
2007.08	467786	103	581	341	211	639	1126	101.0
2007.09	431793	291	717	28	735	745	77	101.0
2007.10	469981	443	261	1	1185	331	3	101.3
2007.11	435640	574	45	0	1423	29	0	101.1
2007.12	950654	748	17	1	1873	27	0	101.3
2008.01	438998	302	4	2	835	5	0	101.1
2008.02	476282	309	4	0	884	5	0	100.9

2008.03	453482	291	5	4	905	6	0	101.4
2008.04	469774	232	12	28	676	18	43	101.3
2008.05	435076	192	30	148	471	39	293	102.1
2008.06	737166	150	102	222	358	95	661	102.6
2008.07	587732	103	236	400	227	245	1212	102.8
2008.08	488216	88	615	307	197	670	1012	103.1
2008.09	433502	278	625	28	827	693	125	103.1
2008.10	481746	445	241	2	1336	337	7	103.0
2008.11	439394	526	36	0	1601	39	0	102.1
2008.12	969449	661	10	1	1949	13	2	101.7
2009.01	443337	268	5	0	865	17	0	101.1
2009.02	464665	277	3	1	1084	3	0	100.8
2009.03	443429	265	5	0	861	6	2	101.1
2009.04	473779	210	15	32	648	15	56	101.2
2009.05	436123	167	33	141	478	31	301	101.0
2009.06	700239	129	110	243	351	97	735	100.8
2009.07	573821	84	209	329	219	232	1253	100.5
2009.08	466393	80	493	303	193	494	1054	100.8
2009.09	422120	259	522	27	774	686	80	100.8
2009.10	459704	366	204	3	1129	248	5	100.4
2009.11	428219	558	41	1	1732	48	3	100.2
2009.12	906884	525	16	2	1561	17	2	100.0
2010.01	434344	256	7	0	804	5	0	100.1
2010.02	464866	265	2	0	917	3	0	100.0
2010.03	439410	264	5	4	829	8	10	100.3
2010.04	474616	208	12	11	578	21	19	100.4
2010.05	421413	167	31	102	391	31	219	100.3
2010.06	733886	129	96	205	285	93	513	100.1
2010.07	562094	78	183	339	168	161	1054	99.5

2010.08	470717	67	543	327	141	566	935	99.7
2010.09	425771	245	608	22	567	624	36	99.9
2010.10	494398	371	237	2	955	271	5	100.2
2010.11	431281	541	44	1	1538	47	3	99.9
2010.12	895511	533	17	1	1511	23	0	99.6
2011.01	419728	239	6	0	666	6	0	99.5
2011.02	470071	257	6	0	732	6	0	99.5
2011.03	419862	250	8	0	758	13	0	99.8
2011.04	454433	210	16	19	634	27	25	99.9
2011.05	413506	177	37	115	508	54	281	99.9
2011.06	687212	158	84	206	416	70	606	99.7
2011.07	572662	97	162	351	257	138	849	99.7
2011.08	463760	94	487	285	204	508	909	99.9
2011.09	422720	230	453	35	621	517	136	99.9
2011.10	479749	350	215	3	932	220	0	100.0
2011.11	424272	410	41	1	1105	43	2	99.4
2011.12	893811	546	51	0	1487	67	0	99.4
2012.01	430477	252	7	0	574	12	0	99.6
2012.02	483625	268	7	0	647	8	0	99.8
2012.03	441015	257	16	2	505	21	1	100.3
2012.04	469381	199	25	19	355	42	30	100.4
2012.05	417723	158	38	99	312	51	145	100.1
2012.06	712592	129	90	181	208	87	567	99.6
2012.07	557032	97	166	326	179	129	1279	99.3
2012.08	470470	74	519	307	166	503	1087	99.4
2012.09	422046	211	491	44	513	567	118	99.6
2012.10	482101	355	295	2	945	342	3	99.6
2012.11	432681	482	50	1	1572	72	1	99.2
2012.12	902928	508	21	1	1404	29	0	99.3

2013.01	433858	264	8	0	753	13	0	99.3
2013.02	476256	264	9	1	743	13	0	99.2
2013.03	444379	276	16	1	781	22	1	99.4
2013.04	479854	229	30	17	643	42	28	99.7
2013.05	422724	168	41	113	454	58	250	99.8
2013.06	728678	136	82	205	307	99	634	99.8
2013.07	569174	99	169	370	218	154	1204	100.0
2013.08	471411	75	480	284	182	506	862	100.3
2013.09	431931	217	566	27	544	621	85	100.6
2013.10	482684	374	231	2	1045	294	9	100.7
2013.11	436293	417	47	1	1080	56	0	100.8
2013.12	905822	574	25	0	1377	31	1	100.9
2014.01	438646	270	8	2	674	11	6	100.7
2014.02	479268	278	7	0	734	15	0	100.7
2014.03	438145	256	15	1	655	21	4	101.0
2014.04	463964	216	35	20	536	43	27	103.1
2014.05	421117	177	46	114	355	52	273	103.5
2014.06	710375	135	86	190	267	68	453	103.4
2014.07	555276	89	180	315	163	155	1067	103.4
2014.08	463810	82	511	224	147	431	704	103.7
2014.09	421809	236	528	34	574	551	147	103.9
2014.10	488273	379	250	2	1042	247	2	103.6
2014.11	431543	504	57	0	1397	57	1	103.2
2014.12	924911	692	22	0	1555	28	1	103.3
2015.01	440226	301	8	0	847	11	0	103.1
2015.02	488519	307	9	1	820	7	0	102.9
2015.03	449243	327	21	2	842	22	6	103.3
2015.04	476880	262	49	23	604	64	38	103.7
2015.05	430325	186	54	99	364	63	180	104.0

```
2015.06  733589 140   96 203  235   80  529 103.8
2015.07  587156 101  189 297  158  146 1124 103.7
2015.08  475369 103  548 279  212  520  889 103.9
2015.09  415467 272  599  41  655  606  159 103.9
2015.10  485330 397  246   4 1107  252   19 103.9


-------------------
. gen t=_n                        <--- Make data t=1,2,...,190.

. tsset t
        time variable:  t, 1 to 190
                delta:  1 unit

. gen ry=log(y/(cpi/100))         <--- log of real income

. gen rap=log((a/ag)/(cpi/100))  <--- log of real price of apple (yen/1g)

. gen rgp=log((g/gg)/(cpi/100))  <--- log of real price of grape (yen/1g)

. gen rwp=log((w/wg)/(cpi/100))  <--- log of real price of watermelon (yen/1g)
(40 missing values generated)

. gen lwg=log(wg)                 <--- log of demand of watermelon (1g)
(35 missing values generated)

. reg lwg ry rwp rap rgp if lwg>log(10)  <--- OLS using the data for wg>10

      Source |       SS       df       MS              Number of obs =     102
```

```
-------------+------------------------------           F(  4,    97) =   40.23
       Model |  138.187919     4  34.5469797           Prob > F      =  0.0000
    Residual |  83.2907458    97  .858667482           R-squared     =  0.6239
-------------+------------------------------           Adj R-squared =  0.6084
       Total |  221.478665   101  2.19285807           Root MSE      =  .92664

------------------------------------------------------------------------------
         lwg |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ry |   .8102279   .5131024     1.58   0.118    -.2081383    1.828594
         rwp |  -1.575157   .3569839    -4.41   0.000    -2.283671   -.8666422
         rap |   2.854632   .6983476     4.09   0.000     1.468606    4.240659
         rgp |   2.158679   .6110691     3.53   0.001     .9458762    3.371482
       _cons |  -3.826122   6.894227    -0.55   0.580    -17.50925    9.857011
------------------------------------------------------------------------------

. truncreg lwg ry rwp rap rgp if lwg>log(10)  <--- This is equivalent to OLS
(note: 0 obs. truncated)

Fitting full model:

Iteration 0:   log likelihood = -134.46068
Iteration 1:   log likelihood = -134.39761
Iteration 2:   log likelihood = -134.39733
Iteration 3:   log likelihood = -134.39733

Truncated regression
Limit:   lower =       -inf                             Number of obs =    102
         upper =       +inf                             Wald chi2(4)  = 169.23
```

```
Log likelihood = -134.39733                               Prob > chi2     = 0.0000

------------------------------------------------------------------------------
         lwg |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ry |   .8102279   .5003683     1.62   0.105     -.170476    1.790932
         rwp |  -1.575157   .3481244    -4.52   0.000    -2.257468   -.8928453
         rap |   2.854632   .6810162     4.19   0.000     1.519865    4.189399
         rgp |   2.158679   .5959037     3.62   0.000     .9907293    3.326629
       _cons |  -3.826122   6.723128    -0.57   0.569    -17.00321    9.350967
-------------+----------------------------------------------------------------
      /sigma |   .9036459   .0632679    14.28   0.000     .7796432    1.027649
------------------------------------------------------------------------------

. truncreg lwg ry rwp rap rgp if lwg>log(10), ll(log(10))   <--- truncated reg
(note: 0 obs. truncated)

Fitting full model:

Iteration 0:   log likelihood = -132.93358
Iteration 1:   log likelihood = -132.70871
Iteration 2:   log likelihood = -132.70789
Iteration 3:   log likelihood = -132.70789

Truncated regression
Limit:   lower =  2.3025851                               Number of obs =     102
         upper =       +inf                               Wald chi2(4)  = 145.68
Log likelihood = -132.70789                               Prob > chi2   = 0.0000
```

lwg	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
ry	.760959	.5179994	1.47	0.142	-.2543011	1.776219
rwp	-1.682078	.3724194	-4.52	0.000	-2.412006	-.952149
rap	2.958551	.7114935	4.16	0.000	1.564049	4.353053
rgp	2.299172	.6349926	3.62	0.000	1.054609	3.543734
_cons	-3.212068	6.960658	-0.46	0.644	-16.85471	10.43057
/sigma	.9260598	.0686138	13.50	0.000	.7915792	1.06054

$$\log(wg_t) = \beta_0 + \beta_1 \log(ry_t) + \beta_2 \log(rwp_t) + \beta_3 \log(rap_t) + \beta_4 \log(rgp_t)$$

Pick up the cases of $wg_t > 10$.

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:

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