Example 5: Multinomial logit model:

The ith individual has $m+1$ choices, i.e., $j = 0, 1, \cdots, m$.

$$P(y_i = j) = \frac{\exp(X_i\beta_j)}{\sum_{j=0}^{m} \exp(X_i\beta_j)} \equiv P_{ij},$$

for $\beta_0 = 0$. The case of $m = 1$ corresponds to the bivariate logit model (binary choice).

Note that

$$\log \frac{P_{ij}}{P_{i0}} = X_i\beta_j$$

The log-likelihood function is:

$$\log L(\beta_1, \cdots, \beta_m) = \sum_{i=1}^{n} \sum_{j=0}^{m} d_{ij} \ln P_{ij},$$

where $d_{ij} = 1$ when the ith individual chooses jth choice, and $d_{ij} = 0$ otherwise.

Example 6: Nested logit model:

(i) In the 1st step, choose YES or NO. Each probability is P_Y and $P_N = 1 - P_Y$.

(ii) Stop if NO is chosen in the 1st step. Go to the next if YES is chosen in the 1st step.

(iii) In the 2nd step, choose A or B if YES is chosen in the 1st step. Each probability is $P_{A|Y}$ and $P_{B|Y}$.

For simplicity, usually we assume the logistic distribution.

So, we call the nested logit model.

The probability that the ith individual chooses NO is:

$$P_{N,i} = \frac{1}{1 + \exp(X_i\beta)}.$$

The probability that the ith individual chooses YES and A is:

$$P_{A|Y,i}P_{Y,i} = P_{A|Y,i}(1 - P_{N,i}) = \frac{\exp(Z_i\alpha)}{1 + \exp(Z_i\alpha)} \frac{\exp(X_i\beta)}{1 + \exp(X_i\beta)}.$$

The probability that the ith individual chooses YES and B is:

$$P_{B|Y,i}P_{Y,i} = (1 - P_{A|Y,i})(1 - P_{N,i}) = \frac{1}{1 + \exp(Z_i\alpha)} \frac{\exp(X_i\beta)}{1 + \exp(X_i\beta)}.$$

In the 1st step, decide if the ith individual buys a car or not.

In the 2nd step, choose A or B.

X_i includes annual income, distance from the nearest station, and so on.

Z_i are speed, fuel-efficiency, car company, color, and so on.

The likelihood function is:

$$\begin{aligned} L(\alpha, \beta) &= \prod_{i=1}^{n} P_{N,i}^{I_{1i}} \Big(((1 - P_{N,i})P_{A|Y,i})^{I_{2i}} ((1 - P_{N,i})(1 - P_{A|Y,i}))^{1-I_{2i}} \Big)^{1-I_{1i}} \\ &= \prod_{i=1}^{n} P_{N,i}^{I_{1i}} (1 - P_{N,i})^{1-I_{1i}} \Big(P_{A|Y,i}^{I_{2i}} (1 - P_{A|Y,i})^{1-I_{2i}} \Big)^{1-I_{1i}}, \end{aligned}$$

where

$$I_{1i} = \begin{cases} 1, & \text{if the } i\text{th individual decides not to buy a car in the 1st step,} \\ 0, & \text{if the } i\text{th individual decides to buy a car in the 1st step,} \end{cases}$$

$$I_{2i} = \begin{cases} 1, & \text{if the } i\text{th individual chooses A in the 2nd step,} \\ 0, & \text{if the } i\text{th individual chooses B in the 2nd step,} \end{cases}$$

Remember that $\mathrm{E}(y_i) = F(X_i\beta^*)$, where $\beta^* = \dfrac{\beta}{\sigma}$.
Therefore, size of β^* does not mean anything.

The marginal effect is given by:

$$\frac{\partial \mathrm{E}(y_i)}{\partial X_i} = f(X_i\beta^*)\beta^*.$$

Thus, the marginal effect depends on the height of the density function $f(X_i\beta^*)$.

7.2 Limited Dependent Variable Model (制限従属変数モデル)

Truncated Regression Model: Consider the following model:

$$y_i = X_i\beta + u_i, \qquad u_i \sim N(0, \sigma^2) \text{ when } y_i > a, \text{ where } a \text{ is a constant,}$$

for $i = 1, 2, \cdots, n$.

Consider the case of $y_i > a$ (i.e., in the case of $y_i \leq a$, y_i is not observed).

$$\mathrm{E}(u_i | X_i\beta + u_i > a) = \int_{a - X_i\beta}^{\infty} u_i \frac{f(u_i)}{1 - F(a - X_i\beta)} \mathrm{d}u_i.$$

Suppose that $u_i \sim N(0, \sigma^2)$, i.e., $\dfrac{u_i}{\sigma} \sim N(0, 1)$.

Using the following standard normal density and distribution functions:

$$\phi(x) = (2\pi)^{-1/2} \exp(-\frac{1}{2}x^2),$$

$$\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp(-\frac{1}{2}z^2)\mathrm{d}z = \int_{-\infty}^{x} \phi(z)\mathrm{d}z,$$

$f(x)$ and $F(x)$ are given by:

$$f(x) = (2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}x^2) = \frac{1}{\sigma}\phi(\frac{x}{\sigma}),$$
$$F(x) = \int_{-\infty}^{x}(2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}z^2)\mathrm{d}z = \Phi(\frac{x}{\sigma}).$$

[Review — Mean of Truncated Normal Random Variable:]

Let X be a normal random variable with mean μ and variance σ^2.

Consider $\mathrm{E}(X|X > a)$, where a is known.

The truncated distribution of X given $X > a$ is:

$$f(x|x > a) = \frac{(2\pi\sigma^2)^{-1/2}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)}{\displaystyle\int_a^{\infty}(2\pi\sigma^2)^{-1/2}\exp\Big(-\frac{1}{2\sigma^2}(x-\mu)^2\Big)\mathrm{d}x} = \frac{\frac{1}{\sigma}\phi(\frac{x}{\sigma})}{1-\Phi(\frac{a-\mu}{\sigma})}.$$

$$\mathrm{E}(X|X>a)=\int_a^\infty xf(x|x>a)\mathrm{d}x=\frac{\displaystyle\int_a^\infty x(2\pi\sigma^2)^{-1/2}\exp\Bigl(-\frac{1}{2\sigma^2}(x-\mu)^2\Bigr)\mathrm{d}x}{\displaystyle\int_a^\infty (2\pi\sigma^2)^{-1/2}\exp\Bigl(-\frac{1}{2\sigma^2}(x-\mu)^2\Bigr)\mathrm{d}x}$$

$$=\frac{\sigma\phi(\frac{a-\mu}{\sigma})+\mu\Bigl(1-\Phi(\frac{a-\mu}{\sigma})\Bigr)}{1-\Phi(\frac{a-\mu}{\sigma})}=\frac{\sigma\phi(\frac{a-\mu}{\sigma})}{1-\Phi(\frac{a-\mu}{\sigma})}+\mu,$$

which are shown below. The denominator is:

$$\begin{aligned}\int_a^\infty (2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\mathrm{d}x &= \int_{\frac{a-\mu}{\sigma}}^\infty (2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z\\ &= 1-\int_{-\infty}^{\frac{a-\mu}{\sigma}} (2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z\\ &= 1-\Phi(\frac{a-\mu}{\sigma}),\end{aligned}$$

where x is transformed into $z=\dfrac{x-\mu}{\sigma}$. $\qquad x>a \implies z=\dfrac{x-\mu}{\sigma}>\dfrac{a-\mu}{\sigma}$.

The numerator is:

$$
\begin{aligned}
&\int_a^{\infty} x(2\pi\sigma^2)^{-1/2}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)\mathrm{d}x \\
&\quad = \int_{\frac{a-\mu}{\sigma}}^{\infty}(\sigma z+\mu)(2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z \\
&\quad = \sigma\int_{\frac{a-\mu}{\sigma}}^{\infty} z(2\pi)^{-1/2}\exp(-\frac{1}{2}z^2)\mathrm{d}z + \mu\int_{\frac{a-\mu}{\sigma}}^{\infty}(2\pi)^{-1/2}\exp(-\frac{1}{2\sigma^2}z^2)\mathrm{d}z \\
&\quad = \sigma\int_{\frac{1}{2}(\frac{a-\mu}{\sigma})^2}^{\infty}(2\pi)^{-1/2}\exp(-t)\mathrm{d}t + \mu\Big(1-\Phi(\frac{a-\mu}{\sigma})\Big) \\
&\quad = \sigma\phi(\frac{a-\mu}{\sigma}) + \mu\Big(1-\Phi(\frac{a-\mu}{\sigma})\Big),
\end{aligned}
$$

where z is transformed into $t=\frac{1}{2}z^2$. $\qquad z>\frac{a-\mu}{\sigma}\implies t=\frac{1}{2}z^2>\frac{1}{2}(\frac{a-\mu}{\sigma})^2$.

[End of Review]

Therefore, the conditional expectation of u_i given $X_i\beta + u_i > a$ is:

$$\mathrm{E}(u_i|X_i\beta + u_i > a) = \int_{a-X_i\beta}^{\infty} u_i \frac{f(u_i)}{1 - F(a - X_i\beta)} \mathrm{d}u_i = \int_{a-X_i\beta}^{\infty} \frac{u_i}{\sigma} \frac{\phi(\dfrac{u_i}{\sigma})}{1 - \Phi(\dfrac{a - X_i\beta}{\sigma})} \mathrm{d}u_i$$

$$= \frac{\sigma\phi(\dfrac{a - X_i\beta}{\sigma})}{1 - \Phi(\dfrac{a - X_i\beta}{\sigma})}.$$

Accordingly, the conditional expectation of y_i given $y_i > a$ is given by:

$$\mathrm{E}(y_i|y_i > a) = \mathrm{E}(y_i|X_i\beta + u_i > a) = \mathrm{E}(X_i\beta + u_i|X_i\beta + u_i > a)$$

$$= X_i\beta + \mathrm{E}(u_i|X_i\beta + u_i > a) = X_i\beta + \frac{\sigma\phi(\dfrac{a - X_i\beta}{\sigma})}{1 - \Phi(\dfrac{a - X_i\beta}{\sigma})},$$

for $i = 1, 2, \cdots, n$.

Estimation:

MLE:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} \frac{f(y_i - X_i\beta)}{1 - F(a - X_i\beta)} = \prod_{i=1}^{n} \frac{1}{\sigma} \frac{\phi(\dfrac{y_i - X_i\beta}{\sigma})}{1 - \Phi(\dfrac{a - X_i\beta}{\sigma})}$$

is maximized with respect to β and σ^2.

Some Examples:

1. Buying a Car:

   $y_i = x_i\beta + u_i$, where y_i denotes expenditure for a car, and x_i includes income, price of the car, etc.

   Data on people who bought a car are observed.

   People who did not buy a car are ignored.

2. Working-hours of Wife:

   y_i represents working-hours of wife, and x_i includes the number of children, age, education, income of husband, etc.

3. Stochastic Frontier Model:

   $y_i = f(K_i, L_i) + u_i$, where y_i denotes production, K_i is stock, and L_i is amount of labor.

   We always have $y_i \leq f(K_i, L_i)$, i.e., $u_i \leq 0$.

   $f(K_i, L_i)$ is a maximum value when we input K_i and L_i.

Censored Regression Model or Tobit Model:

$$y_i = \begin{cases} X_i\beta + u_i, & \text{if } y_i > a, \\ a, & \text{otherwise.} \end{cases}$$

The probability which y_i takes a is given by:

$$P(y_i = a) = P(y_i \le a) = F(a) \equiv \int_{-\infty}^{a} f(x)\mathrm{d}x,$$

where $f(\cdot)$ and $F(\cdot)$ denote the density function and cumulative distribution function of y_i, respectively.

Therefore, the likelihood function is:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} F(a)^{I(y_i=a)} \times f(y_i)^{1-I(y_i=a)},$$

where $I(y_i = a)$ denotes the indicator function which takes one when $y_i = a$ or zero otherwise.

When $u_i \sim N(0, \sigma^2)$, the likelihood function is:

$$L(\beta, \sigma^2) = \prod_{i=1}^{n} \Big(\int_{-\infty}^{a} (2\pi\sigma^2)^{-1/2} \exp(-\frac{1}{2\sigma^2}(y_i - X_i\beta)^2) \mathrm{d}y_i \Big)^{I(y_i=a)}$$
$$\times \Big((2\pi\sigma^2)^{-1/2} \exp(-\frac{1}{2\sigma^2}(y_i - X_i\beta)^2) \Big)^{1-I(y_i=a)},$$

which is maximized with respect to β and σ^2.

Example of Truncated Regression Model:

Demand Function of Watermelon

二人以上の世帯のうち勤労者世帯（2000 年〜）

y	実収入【円】
a	りんご【円】
g	ぶどう【円】
w	すいか【円】
ag	りんご【1g】
gg	ぶどう【1g】
wg	すいか【1g】
CPI	CPI

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 $wg_t > 10$.