ダービン・ワトソン比の定義は次の通りである。

$$DW=\frac{\sum_{i=2}^{n}(\widehat{u}_i-\widehat{u}_{i-1})^2}{\sum_{i=1}^{n}\widehat{u}_i^2}$$

DW は近似的に，次のように表される。

$$\begin{aligned}DW&=\frac{\sum_{i=2}^{n}(\widehat{u}_i-\widehat{u}_{i-1})^2}{\sum_{i=1}^{n}\widehat{u}_i^2}=\frac{\sum_{i=2}^{n}\widehat{u}_i^2-2\sum_{i=2}^{n}\widehat{u}_i\widehat{u}_{i-1}+\sum_{i=2}^{n}\widehat{u}_{i-1}^2}{\sum_{i=1}^{n}\widehat{u}_i^2}\\&=\frac{2\sum_{i=1}^{n}\widehat{u}_i^2-(\widehat{u}_1^2+\widehat{u}_n^2)}{\sum_{i=1}^{n}\widehat{u}_i^2}-2\frac{\sum_{i=2}^{n}\widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=1}^{n}\widehat{u}_i^2}\approx 2(1-\widehat{\rho}),\end{aligned}$$

以下の 2 つの近似が用いられる。

$$\frac{\widehat{u}_1^2+\widehat{u}_n^2}{\sum_{i=1}^{n}\widehat{u}_i^2}\approx 0,$$

$$\frac{\sum_{i=2}^{n}\widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=1}^{n}\widehat{u}_i^2}=\frac{\sum_{i=2}^{n}\widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=2}^{n}\widehat{u}_{i-1}^2+\widehat{u}_n^2}\approx\frac{\sum_{i=2}^{n}\widehat{u}_i\widehat{u}_{i-1}}{\sum_{i=2}^{n}\widehat{u}_{i-1}^2}=\widehat{\rho},$$

すなわち，$\widehat{\rho}$ は $\widehat{u}_i$ と $\widehat{u}_{i-1}$ の回帰係数である。$u_i=\rho u_{i-1}+\epsilon_i$ において，u_i, u_{i-1} の代わりに $\widehat{u}_i, \widehat{u}_{i-1}$ に置き換えて，ρ の推定値 $\widehat{\rho}$ を求める。

1. DW の値が 2 前後のとき，系列相関なし ($\widehat{\rho} = 0$ のとき，$DW \approx 2$)。

2. DW が 2 より十分に小さいとき，正の系列相関と判定される。

3. DW が 2 より十分に大きいとき，負の系列相関と判定される。

正確な判定には，データ数 n とパラメータ数 k に依存する。表 1 を参照せよ。
k' は定数項を除くパラメータ数を表すものとする。

See
`http://www.stanford.edu/~clint/bench/dwcrit.htm`
for the DW table.

Table 1: ダービン・ワトソン統計量の 5 % 点の上限と下限

	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$		$k'=6$		$k'=7$		$k'=8$		$k'=9$		$k'=10$		$k'=11$		$k'=12$		$k'=13$	
n	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du	dl	du
6	0.610	1.400	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.763	1.332	0.559	1.777	0.367	2.287	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.315	2.645	0.203	3.004	—	—	—	—	—	—	—	—	—	—	—	—	—	—
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.380	2.506	0.268	2.832	0.171	3.149	—	—	—	—	—	—	—	—	—	—	—	—
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.444	2.390	0.328	2.692	0.230	2.985	0.147	3.266	—	—	—	—	—	—	—	—	—	—
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360	—	—	—	—	—	—	—	—
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.471	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438	—	—	—	—	—	—
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304	0.098	3.503	—	—	—	—
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184	0.138	3.378	0.087	3.557	—	—
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461	0.407	2.668	0.321	2.873	0.244	3.073	0.177	3.265	0.123	3.441	0.078	3.603
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974	0.220	3.159	0.160	3.335	0.111	3.496
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.691	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885	0.263	3.063	0.200	3.234	0.145	3.395
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.731	2.124	0.637	2.290	0.546	2.461	0.461	2.633	0.380	2.806	0.307	2.976	0.240	3.141	0.182	3.300
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.735	0.349	2.897	0.281	3.057	0.220	3.211
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670	0.391	2.826	0.322	2.979	0.259	3.129
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.750	2.174	0.666	2.318	0.584	2.464	0.506	2.613	0.431	2.761	0.362	2.908	0.297	3.053
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.013	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560	0.470	2.702	0.400	2.844	0.335	2.983
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513	0.508	2.649	0.438	2.784	0.373	2.919
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470	0.544	2.600	0.475	2.730	0.409	2.860
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.959	0.874	2.071	0.798	2.188	0.723	2.309	0.649	2.431	0.578	2.555	0.510	2.680	0.445	2.805
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164	0.753	2.278	0.681	2.396	0.612	2.515	0.544	2.634	0.479	2.754
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141	0.782	2.251	0.712	2.363	0.643	2.477	0.577	2.592	0.513	2.708
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333	0.674	2.443	0.608	2.553	0.545	2.665
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306	0.703	2.411	0.638	2.518	0.576	2.625
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085	0.861	2.181	0.796	2.281	0.731	2.382	0.667	2.484	0.606	2.588
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.079	1.891	1.015	1.978	0.950	2.069	0.885	2.162	0.821	2.257	0.758	2.355	0.695	2.454	0.634	2.553
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236	0.783	2.330	0.722	2.425	0.662	2.521
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.876	1.053	1.957	0.991	2.041	0.930	2.127	0.868	2.216	0.808	2.306	0.748	2.398	0.689	2.492
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948	1.011	2.029	0.951	2.112	0.891	2.197	0.831	2.285	0.772	2.374	0.714	2.464
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939	1.029	2.017	0.970	2.098	0.912	2.180	0.854	2.265	0.796	2.351	0.739	2.438
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932	1.047	2.007	0.990	2.085	0.932	2.164	0.875	2.246	0.819	2.329	0.763	2.413
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.952	2.150	0.896	2.228	0.840	2.309	0.785	2.391
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.022	1.038	2.088	0.988	2.156	0.938	2.225	0.887	2.296
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.986	1.110	2.044	1.064	2.103	1.019	2.163	0.973	2.225
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010	1.129	2.062	1.087	2.116	1.045	2.170
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850	1.298	1.894	1.260	1.939	1.222	1.984	1.184	2.031	1.145	2.079	1.106	2.127
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964	1.231	2.006	1.195	2.049	1.160	2.093
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1.838	1.369	1.874	1.337	1.910	1.305	1.948	1.272	1.987	1.239	2.026	1.206	2.066
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935	1.308	1.970	1.277	2.006	1.247	2.043
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925	1.340	1.957	1.312	1.990	1.283	2.024
85	1.623	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916	1.369	1.946	1.342	1.977	1.315	2.008
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909	1.395	1.937	1.369	1.966	1.344	1.995
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827	1.489	1.852	1.465	1.877	1.442	1.903	1.418	1.930	1.394	1.956	1.370	1.984
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826	1.506	1.850	1.484	1.874	1.462	1.898	1.439	1.923	1.416	1.948	1.393	1.974
150	1.720	1.747	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832	1.622	1.846	1.608	1.862	1.593	1.877	1.579	1.892	1.564	1.908	1.549	1.924
200	1.758	1.779	1.748	1.789	1.738	1.799	1.728	1.809	1.718	1.820	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874	1.654	1.885	1.643	1.897	1.632	1.908

n は標本数，k' は定数項を除く説明変数の数とする。

（出所）`http://www.stanford.edu/~clint/bench/dwcrit.htm`

$$DW = \frac{\sum_{i=2}^{n}(\widehat{u_i} - \widehat{u}_{i-1})^2}{\sum_{i=1}^{n}\widehat{u_i^2}} \approx 2(1-\widehat{\rho}) \longrightarrow 2(1-\rho)$$

$-1 < \rho < 1$ なので (証明略)，近似的に $0 \le DW \le 4$ となる。

- $0 \le DW \le dl \qquad \longrightarrow$ u_i に正の系列相関
- $dl \le DW \le du \qquad \longrightarrow$ u_i に正の系列相関と判定できない
- $du \le DW \le 4 - du \qquad \longrightarrow$ u_i に系列相関なし
- $4 - du \le DW \le 4 - dl \qquad \longrightarrow$ u_i に負の系列相関と判定できない
- $4 - dl \le DW \le 4 \qquad \longrightarrow$ u_i に負の系列相関

数値例：　今までと同じ数値例で，DW を計算する。

i	Y_i	X_i	X_iY_i	X_i^2	$\widehat{Y}_i$	$\widehat{u}_i$
1	6	10	60	100	6.8	−0.8
2	9	12	108	144	8.1	0.9
3	10	14	140	196	9.4	0.6
4	10	16	160	256	10.7	−0.7
合計	$\sum Y_i$	$\sum X_i$	$\sum X_iY_i$	$\sum X_i^2$	$\sum \widehat{Y}_i$	$\sum \widehat{u}_i$
	35	52	468	696	35	0
平均	$\overline{Y}$	$\overline{X}$				
	8.75	13				

$$
\begin{aligned}
DW &= \frac{\sum_{i=2}^{n}(\widehat{u}_i - \widehat{u}_{i-1})^2}{\sum_{i=1}^{n}\widehat{u}_i^2} \\
&= \frac{(-0.8-0.9)^2 + (0.9-0.6)^2 + (0.6-(-0.7))^2}{(-0.8)^2 + 0.9^2 + 0.6^2 + (-0.7)^2} = \frac{4.67}{2.30} = 2.03
\end{aligned}
$$

推定結果の表記方法：　回帰モデル：

$$Y_i = \alpha + \beta X_i + u_i,$$

の推定の結果，$\widehat{\alpha} = 0.3, \widehat{\beta} = 0.65, s_{\widehat{\alpha}} = \sqrt{10.0005} = 3.163, s_{\widehat{\beta}} = \sqrt{0.0575} = 0.240,$
$\dfrac{\widehat{\alpha}}{s_{\widehat{\alpha}}} = 0.095, \dfrac{\widehat{\beta}}{s_{\widehat{\beta}}} = 2.708, s^2 = 1.15$ (すなわち，$s = 1.07$), $R^2 = 0.786, \overline{R}^2 = 0.679,$
$DW = 2.03$ を得た。

これらをまとめて，

$$Y_i = \underset{(0.095)}{0.3} + \underset{(2.708)}{0.65} X_i,$$

$$R^2 = 0.786, \quad \overline{R}^2 = 0.679, \quad s = 1.07, \quad DW = 2.03,$$

ただし，係数の推定値の下の括弧内は t 値を表すものとする。

または，

$$
Y_i = \underset{(3.163)}{0.3} + \underset{(0.240)}{0.65} X_i,
$$

$$
R^2 = 0.786, \quad \overline{R}^2 = 0.679, \quad s = 1.07, \quad DW = 2.03,
$$

ただし，係数の推定値の下の括弧内は標準誤差を表すものとする。

のように書く。$s = \sqrt{1.15} = 1.07$ に注意。

4.2　系列相関のもとで回帰式の推定

回帰式が

$$
Y_i = \alpha + \beta X_i + u_i,
$$

$$
u_i = \rho u_{i-1} + \epsilon_i,
$$

のときの推定を考える。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とする。

u_i を消去すると，

$$(Y_i - \rho Y_{i-1}) = \alpha(1-\rho) + \beta(X_i - \rho X_{i-1}) + \epsilon_i,$$

となり，

$Y_i^* = (Y_i - \rho Y_{i-1}),\ X_i^* = (X_i - \rho X_{i-1})$

を新たな変数として，

$$Y_i^* = \alpha' + \beta X_i^* + \epsilon_i,$$

に最小二乗法を適用する。$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とするなので，最小二乗法を適用が可能となる。ただし，$\alpha' = \alpha(1-\rho)$ の関係が成り立つことに注意。

より一般的に，回帰式が

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i,$$

$$u_i = \rho u_{i-1} + \epsilon_i,$$

のときの推定を考える。ただし，$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とする。

u_i を消去すると，

$$(Y_i - \rho Y_{i-1}) = \beta_1(X_{1i} - \rho X_{1,i-1}) + \beta_2(X_{1i} - \rho X_{2,i-1}) + \cdots + \beta_k(X_{1i} - \rho X_{k,i-1}) + \epsilon_i,$$

となり，

$Y_i^* = (Y_i - \rho Y_{i-1}),\ X_{1i}^* = (X_{1i} - \rho X_{1,i-1}),\ X_{2i}^* = (X_{2i} - \rho X_{2,i-1}),\ \cdots,\ X_{ki}^* = (X_{ki} - \rho X_{k,i-1})$ を新たな変数として，

$$Y_i^* = \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + \cdots + \beta_k X_{ki}^* + \epsilon_i$$

最小二乗法を適用する。$\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ は互いに独立とするなので，最小二乗法を適用が可能となる。

ρ の求め方について (その 1):　DW は近似的に $DW \approx 2(1 - \widehat{\rho})$ と表されるので，DW から ρ の推定値 $\widehat{\rho}$ を逆算して，

$Y_i^* = (Y_i - \widehat{\rho} Y_{i-1})$, $X_{1i}^* = (X_{1i} - \widehat{\rho} X_{1,i-1})$, $X_{2i}^* = (X_{2i} - \widehat{\rho} X_{2,i-1})$, $\cdots$, $X_{ki}^* = (X_{ki} - \widehat{\rho} X_{k,i-1})$ を新たな変数として，

$$Y_i^* = \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + \cdots + \beta_k X_{ki}^* + \epsilon_i,$$

に最小二乗法を適用する。

ρ の求め方について (その 2):　収束計算によって求める。$\longrightarrow$ コクラン・オーカット法

1. $Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i, \quad i = 1, 2, \cdots, n$

   を最小二乗法で推定する。$\longrightarrow$ $\widehat{\beta}_1, \cdots, \widehat{\beta}_k, \widehat{u}_i$ を得る。

2. $\widehat{u}_i = \rho \widehat{u}_{i-1} + \epsilon_i, \quad i = 2, 3, \cdots, n$

   を最小二乗法で推定する。$\longrightarrow$ $\widehat{\rho}$ を得る。

3. $\rho^{(m-1)} = \widehat{\rho}$ とおく。

4. $Y_i^* = (Y_i - \rho^{(m-1)} Y_{i-1}),\ X_{1i}^* = (X_{1i} - \rho^{(m-1)} X_{1,i-1}),\ X_{2i}^* = (X_{2i} - \rho^{(m-1)} X_{2,i-1}),\ \cdots,$

   $X_{ki}^* = (X_{ki} - \rho^{(m-1)} X_{k,i-1})$ を計算する。

   $$Y_i^* = \beta_1 X_{1i}^* + \beta_2 X_{2i}^* + \cdots + \beta_k X_{ki}^* + \epsilon_i, \quad i = 2, 3, \cdots, n$$

   を最小二乗法で推定する。 $\longrightarrow \widehat{\beta}_1, \cdots, \widehat{\beta}_k$ を得る。

5. $\widehat{u}_i = Y_i - \widehat{\beta}_1 X_{1i} - \widehat{\beta}_2 X_{2i} - \cdots - \widehat{\beta}_k X_{ki}, \quad i = 1, 2, \cdots, n$

   を計算する。

6. ステップ 2 に戻り， $m = 1, 2, \cdots$ について繰り返す。

   収束先を $\beta_1, \beta_2, \cdots, \beta_k, \rho$ の推定値とする。

5　不均一分散(不等分散)

回帰式が

$$Y_i = \alpha + \beta X_i + u_i$$

の場合を考える。X_i が外生変数，Y_i は内生変数，u_i は互いに独立な同一の分布を持つ攪乱項(最小二乗法に必要な仮定)とする。「独立な同一の分布」の意味は「攪乱項 $u_1, u_2, \cdots, u_n$ はそれぞれ独立に平均ゼロ，分散 σ^2 の分布する」である。分散が時点に依存する場合，代表的には，分散が他の変数(例えば，z_i)に依存する場合，すなわち，u_i の平均はゼロ，分散は $\sigma_*^2 z_i^2$ の場合は，最小二乗法の仮定に反する。そのため，単純には，$Y_i = \alpha + \beta X_i + u_i$ に最小二乗法を適用できない。以下のような修正が必要となる。

$$\frac{Y_i}{z_i} = \alpha \frac{1}{z_i} + \beta \frac{X_i}{z_i} + \frac{u_i}{z_i} = \alpha \frac{1}{z_i} + \beta \frac{X_i}{z_i} + u_i^*$$

このとき，新たな攪乱項 u_i^* は平均ゼロ，分散 σ_*^2 の分布となる(すなわち,「同

一の」分布)。

$$\mathrm{E}(u_i^*) = \mathrm{E}\left(\frac{u_i}{z_i}\right) = \left(\frac{1}{z_i}\right)\mathrm{E}(u_i) = 0$$

u_i の仮定 $\mathrm{E}(u_i) = 0$ が使われている。

$$\mathrm{V}(u_i^*) = \mathrm{V}\left(\frac{u_i}{z_i}\right) = \left(\frac{1}{z_i}\right)^2 \mathrm{V}(u_i) = \sigma_*^2$$

u_i の仮定 $\mathrm{V}(u_i) = \sigma_*^2 z_i^2$ が最後に使われている。
よって，$\dfrac{Y_i}{z_i}, \dfrac{1}{z_i}, \dfrac{X_i}{z_i}$ を新たな変数として，最小二乗法を適用することができる。

不均一分散の検定について

$$\widehat{u_i^2} = \gamma z_i + \epsilon_i$$

を推定し，γ の推定値 $\widehat{\gamma}$ の有意性の検定を行う (通常の t 検定)。

z_i は回帰式に含まれる変数でもよい。例えば，u_i の平均はゼロ，分散は $\sigma_*^2 X_i^2$ の

場合，各変数を X_i で割って，

$$\frac{Y_i}{X_i} = \alpha \frac{1}{X_i} + \beta + \frac{u_i}{X_i} = \alpha \frac{1}{X_i} + \beta + u_i^*$$

を推定すればよい。β は定数項として推定されるが，意味は限界係数 (すなわち，傾き) と同じなので注意すること。

6　最尤法について

標本 $X_1, X_2, \cdots, X_n$ の密度関数：

$$f(x_1, x_2, \cdots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

θ は未知母数 $\Longrightarrow \widehat{\theta}_n(x_1, x_2, \cdots, x_n)$ によって推定

$$l(\theta) = l(\theta; x) = l(\theta; x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(x_i; \theta)$$

のように，θ の関数として考える。

$l(\theta)$： 尤度関数

尤度関数を最大にする θ を $\widehat{\theta}_n$ とする。

$\widehat{\theta}_n \equiv \widehat{\theta}_n(X_1, X_2, \cdots, X_n) \Longrightarrow$ 最尤推定量

$\widehat{\theta}_n(x_1, x_2, \cdots, x_n) \Longrightarrow$ 最尤推定値

すなわち，

$$\frac{\partial l(\theta)}{\partial \theta} = 0$$

を解くことによって，最尤推定量 $\widehat{\theta}_n \equiv \widehat{\theta}_n(X_1, X_2, \cdots, X_n)$ が得られる。

最尤推定量の性質：

小標本について (n が小さいとき)：

- 一般に，最尤推定量は不偏性を持っていないが，適当な変換によって，不偏推定量を作ることが出来る場合が多い。

- 有効推定量が存在すれば (すなわち，クラメール・ラオの不等式の等号を満たすような推定量が存在するならば)，最尤推定量は有効推定量に一致する。

- 十分統計量が存在すれば，最尤推定量は十分統計量の関数となる。

大標本について (n が大きいとき)：

$n \longrightarrow \infty$ のとき，

$$\sqrt{n}(\widehat{\theta}_n - \theta) \longrightarrow N\Big(0, \sigma^2\Big)$$

となる。$\Longrightarrow$ 一致性，漸近的正規性，漸近的有効性

ただし，

$$\sigma^2 = \sigma^2(\theta) = \frac{1}{\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\theta)}{\partial\theta}\right)^2\right]}$$

すなわち，$n \longrightarrow \infty$ のとき，

$$\frac{\sqrt{n}(\widehat{\theta}_n - \theta)}{\sigma(\theta)} = \frac{\widehat{\theta}_n - \theta}{\sigma(\theta)/\sqrt{n}} \longrightarrow N(0, 1)$$

となる。

したがって，厳密ではないが，n が大きいとき，

$$\widehat{\theta}_n \sim N\left(\theta, \frac{\sigma^2(\theta)}{n}\right)$$

と近似できる。

すなわち，$n \longrightarrow \infty$ のとき，$\widehat{\theta}_n$ の分散はクラメール・ラオの不等式の下限 $\dfrac{\sigma^2(\theta)}{n}$ に近づくことを意味する。

$\Longrightarrow$ 漸近的に有効推定量

さらに，分母の θ を最尤推定量 $\widehat{\theta}_n$ で置き換えて，$n \longrightarrow \infty$ のとき，

$$\frac{\widehat{\theta}_n - \theta}{\sigma(\widehat{\theta}_n)/\sqrt{n}} \longrightarrow N(0, 1)$$

となる。

実際には，n が大きいとき，

$$\widehat{\theta}_n \sim N\left(\theta, \frac{\sigma^2(\widehat{\theta}_n)}{n}\right)$$

と近似して用いる。

例：

$X_1, X_2, \cdots, X_n$ は互いに独立で，すべて同一の正規分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。μ, σ^2 の最尤推定量を求める。

$$f(x_1, x_2, \cdots, x_n; \mu, \sigma^2) = \prod_{i=1}^{n} f(x_i; \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right)$$
$$= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right) = l(\mu, \sigma^2)$$

対数をとる。(最大化しやすくなる場合が多い)

$$\log l(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

対数尤度関数 $\log l(\mu, \sigma^2)$ を μ と σ^2 について微分して，ゼロと置く。

$$\frac{\partial \log l(\mu, \sigma^2)}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0$$

$$\frac{\partial \log l(\mu, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0$$

この 2 つの連立方程式を解く。

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

μ, σ^2 の最尤推定量は,

$$\overline{X}, \quad S^{**2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

となる。

$\mathrm{E}(\overline{X}) = \mu$ なので，μ の最尤推定量 $\overline{X}$ は不偏推定量である。

$\mathrm{E}(S^{**2}) = \dfrac{n-1}{n}\sigma^2 \neq \sigma^2$ なので，σ^2 の最尤推定量 S^{**2} は不偏推定量でない。

例：

$X_1, X_2, \cdots, X_n$ は互いに独立で，すべて同一のベルヌイ分布ですべて同一の分布)に従うものとする。すなわち，X の確率関数は $P(X = x) = f(x; p) = p^x(1-p)^{1-p}$, $x = 0, 1,$ となる。p の最尤推定量を求める。

$$f(x_1, x_2, \cdots, x_n; p) = \prod_{i=1}^{n} f(x_i; p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_i x_i}(1-p)^{n-\sum_i x_i} = l(p)$$

対数をとる。

$$\log l(p) = (\sum_i x_i)\log(p) + (n - \sum_i x_i)\log(1-p)$$

対数尤度関数 $\log l(p)$ を p について微分して，ゼロと置く。

$$\frac{\partial \log l(p)}{\partial p} = \frac{\sum_i x_i}{p} - \frac{n - \sum_i x_i}{1-p} = \frac{\sum_i x_i - np}{p(1-p)} = 0$$

この方程式を解く。

$$p = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

p の最尤推定量は,

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

となる。

$\mathrm{E}(\overline{X}) = p$ なので，p の最尤推定量 $\overline{X}$ は不偏推定量である。

X がベルヌイ分布 $f(x; p)$ のとき，$\mathrm{E}(X) = p$ に注意。