6. Quasi-likelihood and Estimating Equations

§6.1. Quasi-likelihood

- A review of IWLS for GLIM
  
  The IWLS procedure for a GLIM iteratively solve
  
  \[ [X^tWX]\beta = X^tWz, \]
  
  where
  
  \[ W = \text{Diag}([\frac{\partial \eta_i}{\partial \mu_i}]^{-2}[V(\mu_i)]^{-1}), \]
  
  \[ z_i = \eta_i + \frac{\partial \eta_i}{\partial \mu_i}(y_i - \mu_i). \]

  The necessary component required for IWLS:

  Variance function: \( V(\mu). \)
  
  Link function: \( \eta = g(\mu). \)

- Quasi-likelihood model
  
  Data:  \((y_i, x_i), i = 1, \ldots, n.\)
  
  Assumption:
  
  \[ y_i \sim (\mu_i, \sigma^2 V(\mu_i)), \]
  
  \[ \eta_i = g(\mu_i), \]
  
  \[ \eta_i = x_i^t \beta. \]
Quasi-likelihood:
For a random variable $Y$ with mean $\mu$ and variance $\sigma^2 V(\mu)$, define

$$Q(\mu, y) = \int_{y}^{\mu} \frac{y - s}{\sigma^2 V(s)} ds.$$  

If the defining integral exists, $Q(\mu, y)$ is called the log quasi-likelihood function of $\mu$ based on $Y$.

The rationale of the Quasi-likelihood: Consider

$$\frac{\partial Q(\mu, y)}{\partial \mu} = \frac{y - \mu}{\sigma^2 V(\mu)}.$$

Then

$$E \frac{\partial Q(\mu, y)}{\partial \mu} = 0,$$

$$\text{Var} \left( \frac{\partial Q(\mu, y)}{\partial \mu} \right) = \frac{1}{\sigma^2 V(\mu)} = -E \frac{\partial^2 Q(\mu, y)}{\partial \mu^2}.$$

The properties of the first two derivatives of $Q(\mu, y)$ are the same as those of a log likelihood function. Hence $Q(\mu, y)$ can be considered as the log of some probability density function $f(\mu, y)$.

(a) $Q(\mu, y)$ differs from $\ln f(\mu, y)$ possibly by a quantity which does not involve $\mu$.
(b) $Q(\mu, y)$ is uniquely determined by $V(\mu)$. 


Typical $V(\mu)$ and corresponding $Q(\mu, y)$

<table>
<thead>
<tr>
<th>$V(\mu)$</th>
<th>$Q(\mu, y)$</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$-(y - \mu)^2/2$</td>
<td>Normal</td>
</tr>
<tr>
<td>$\mu$</td>
<td>$y \ln \mu - \mu$</td>
<td>Poisson</td>
</tr>
<tr>
<td>$\mu^2$</td>
<td>$-y/\mu - \ln \mu$</td>
<td>Gamma</td>
</tr>
<tr>
<td>$\mu^3$</td>
<td>$-y/(2\mu^2) + 1/\mu$</td>
<td>Inverse Gaussian</td>
</tr>
<tr>
<td>$\mu^\zeta$</td>
<td>$\mu^{-\zeta}(\frac{\mu y}{1-\zeta} - \frac{\mu^2}{2-\zeta})$</td>
<td>—</td>
</tr>
<tr>
<td>$\mu(1-\mu)$</td>
<td>$y \ln \frac{\mu}{1-\mu} + \ln(1-\mu)$</td>
<td>Binomial/m</td>
</tr>
<tr>
<td>$\mu^2(1-\mu)^2$</td>
<td>$(2y-1) \ln \frac{\mu}{1-\mu} - \frac{y}{\mu} - \frac{1-y}{1-\mu}$</td>
<td>—</td>
</tr>
<tr>
<td>$\mu + \mu^2/k$</td>
<td>$y \ln \frac{\mu}{k+\mu} + k \ln \frac{k}{k+\mu}$</td>
<td>Negative Binomial</td>
</tr>
</tbody>
</table>

Note: A dispersion parameter $\phi$ can be incorporated into each of the $Q(\mu, y)$ as $Q(\mu, y)/\phi$.

- **Estimation and inference on Quasi-likelihood model**

The parameter $\beta$ is estimated by maximizing the log quasi-likelihood function:

$$Q(\mu(\beta), y) = \sum_{i=1}^{n} Q(\mu_i, y_i),$$

where

$$\mu_i = g^{-1}(\eta_i), \quad \eta_i = x_i^t \beta.$$ 

The maximization is carried out by solving

$$u(\beta) = \frac{\partial Q(\mu(\beta), y)}{\partial \beta} = 0.$$
The \( \mathbf{u}(\mathbf{\beta}) \) is called the vector of quasi-scores. It can be obtained that

\[
\mathbf{u}(\mathbf{\beta}) = X^t \mathbf{W} \frac{\partial \eta}{\partial \mu^t} (\mathbf{y} - \mu)/\sigma^2,
\]

\[
A(\mathbf{\beta}) = -E \frac{\partial^2 Q(\mu(\mathbf{\beta}), \mathbf{y})}{\partial \mathbf{\beta} \partial \mathbf{\beta}^t} = X^t \mathbf{W} \mathbf{X}/\sigma^2,
\]

where \( \mathbf{W} = \text{Diag}(\frac{\partial \eta_i}{\partial \mu_i})^{-2}[V(\mu_i)]^{-1}) \).

**The IWLS procedure**

The Newton-Rhapson procedure with Fisher scoring is again equivalent to an IWLS which solves iteratively

\[
X^t \mathbf{W} (\mathbf{\beta}^{\text{OLD}}) \mathbf{X} \mathbf{\beta}^{\text{NEW}} = X^t \mathbf{W} (\mathbf{\beta}^{\text{OLD}}) \mathbf{z}(\mathbf{\beta}^{\text{OLD}}),
\]

where \( \mathbf{z} = \eta + \frac{\partial \eta}{\partial \mu^t} (\mathbf{y} - \mu) \).

Another expression for \( \mathbf{u}(\mathbf{\beta}) \) and \( A(\mathbf{\beta}) \):

\[
\mathbf{u}(\mathbf{\beta}) = D^t V^{-1} (\mathbf{y} - \mu)/\sigma^2,
\]

\[
A(\mathbf{\beta}) = D^t V^{-1} D/\sigma^2,
\]

where \( D = \frac{\partial \mu}{\partial \mathbf{\beta}^t} \).

**Estimation of \( \sigma^2 \):**

\[
\hat{\sigma}^2 = \frac{1}{n - p} \sum_{i=1}^{n} (y_i - \hat{\mu}_i)^2 / V(\hat{\mu}_i).
\]
The asymptotic distribution of $\hat{\beta}$:

$$\hat{\beta} \sim N(\beta, \sigma^2[D^tV^{-1}D]^{-1}).$$

The asymptotic distribution provides the basis for the construction of confidence intervals and hypothesis testing. The Wald tests can be constructed from the asymptotic distribution.

§6.2*. Extended Quasi-likelihood

- The first order extended quasi-likelihood

Let $Y \sim (\mu, \sigma^2V(\mu))$. Suppose that $\sigma^2$ is small and $\kappa_{r+1} = \kappa'_r\kappa_2$ where $\kappa_j$ is the $j$th cumulant of $Y$ and $\kappa'_r$ is the derivative of $\kappa_r$ with respect to $\mu$. The first order extended quasi likelihood for $\mu$ and $\sigma^2$ is given by

$$Q^+(\mu, \sigma^2, y) = -\frac{D(y, \mu)}{2\sigma^2} - \frac{1}{2} \ln \sigma^2,$$

where $D(y, \mu) = 2 \int_{\mu}^{y} \frac{y-s}{V(s)} ds$.

Under the above assumption,

$$ED(\mu, y) \approx \sigma^2, \quad E\frac{\partial Q^+(\mu, \sigma^2, y)}{\partial \mu} = 0, \quad E\frac{\partial Q^+(\mu, \sigma^2, y)}{\partial \sigma^2} \approx 0.$$

and, further,

$$I(\mu, \sigma^2) \approx \begin{pmatrix} \frac{1}{\sigma^2V(\mu)} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}.$$
**The construction of extended quasi likelihood**

It is desired to construct $Q^+(\mu, \sigma^2, y)$ such that

$$E\frac{\partial Q^+(\mu, \sigma^2, y)}{\partial \mu} = 0, \quad E\frac{\partial Q^+(\mu, \sigma^2, y)}{\partial \sigma^2} = 0.$$ 

Motivated by the form of exponential family, let

$$Q^+(\mu, \sigma^2, y) = Q(\mu, y) + h(\sigma^2, y) = \frac{D(y, \mu)}{2\sigma^2} + h(\sigma^2, y).$$

Let

$$h(\sigma^2, y) = -\frac{1}{2}h_1(\sigma^2) - h_2(y).$$

Thus,

$$E\frac{\partial Q^+(\mu, \sigma^2, y)}{\partial \sigma^2} = 0 \iff \frac{ED(y, \mu)}{\sigma^4} - \frac{h'_1(\sigma^2)}{\sigma^4} = 0 \iff \sigma^4 h'_1(\sigma^2) = ED(y, \mu).$$

Expand $1/V(s)$ at $s = \mu$,

$$\frac{1}{V(s)} = \frac{1}{V(\mu)} - \frac{V'}{V^2}(t - \mu) + \frac{1}{2}\left[\frac{2(V')^2}{V^3} - \frac{V V''}{V^3}\right](t - \mu)^2 + \cdots$$
Note

\[ \int_{\mu}^{y} (s - \mu)(y - s) ds = \frac{1}{6}(y - \mu)^3, \]
\[ \int_{\mu}^{y} (s - \mu)^2(y - s) ds = \frac{1}{12}(y - \mu)^4. \]

Hence

\[ ED(y, \mu) \approx \sigma^2 - \frac{V'}{3V^2} E(Y - \mu)^3 + \frac{1}{12}\frac{2(V')^2 - VV''}{V^3} E(Y - \mu)^4 \]
\[ = \sigma^2 - \frac{V'}{3V^2}\kappa_3 + \frac{1}{12}\frac{2(V')^2 - VV''}{V^3}(\kappa_4 + 3\kappa_2^2) \]
\[ = \sigma^2 + \frac{1}{12V^2}\left[6\sigma^4 V(V')^2 - 3\sigma^4 V^2 V'' - 4V'\kappa_3\right] \]
\[ + \frac{\kappa_4}{12}\frac{2(V')^2 - VV''}{V^3}. \]

Since \( \kappa_{r+1} = \kappa'_r\kappa_2 \),

\( \kappa_3 = \sigma^4 VV', \quad \kappa_4 = \sigma^6 [VV']' V = O(\sigma^6). \)

Thus, if \( \sigma^2 \) is small,

\[ ED(y, \mu) \approx \sigma^2, \]
— first order approximation,
\[ ED(y, \mu) \approx \sigma^2 + \frac{\sigma^4}{12}\left[2(V')^2/V - 3V''\right], \]
— second order approximation.
The first order approximation to \( h_1(\sigma^2) \):

\[
h_1(\sigma^2) = \ln \sigma^2.
\]

The second order approximation to \( h_1(\sigma^2) \):

\[
h_1(\sigma^2) = \ln \sigma^2 + \frac{\sigma^2}{12} [2(V')^2 / V - 3V''].
\]

It can also be obtained by \( \delta \)-method that

\[
\text{Var}(D(Y, \mu)) \approx 2\sigma^4,
\]

\[
\text{Cov}(D(Y, \mu), Y) \approx (\kappa_3 - \kappa_2\kappa'_2) / V = 0,
\]

which gives rise to the information matrix \( I(\mu, \sigma^2) \).

\section*{§6.3. Estimating equations}

- **Estimating functions**

Let \( y \) be the data vector and \( \theta \) the parameter to be estimated. An estimating function is a function \( g(y, \theta) \) of \( y \) and \( \theta \) such that

\[
E_\theta[g(y, \theta)] = 0, \quad \text{for all } \theta.
\]

Remark: The quasi score function \( u(y, \beta) \) is an estimating function, since

\[
 u(y, \beta) = D^t V^{-1}(y - \mu(\beta)) / \sigma^2,
\]

\[
 Eu(y, \beta) = 0.
\]
• **Estimating equation estimate**
  Suppose \( g(y, \theta) \) is an estimating function. Then
  \[ g(y, \theta) = 0 \]
  is called an estimating equation. The solution \( \hat{\theta} \) of the estimating equation (EE) is called an EE estimate.

• **The optimal property of the quasi score function**
  Among all linear estimating functions of the form
  \[
h(y, \beta) = H^t(y - \mu(\beta)),
\]
  where \( H \) is a \( n \times p \) matrix which might be a function of \( \beta \) but not of \( y \), the quasi score function \( u(y, \beta) \) is optimal in the sense that the EE estimate of \( \beta \) based on \( u(y, \beta) \) has the minimum variance; that is, if \( \hat{\beta} \) and \( \tilde{\beta} \) are EE estimates based on \( u(y, \beta) \) and \( h(y, \beta) \) respectively, and \( a \) is any constant vector, then
  \[
  \text{Var}(a^t \hat{\beta}) \leq \text{Var}(a^t \tilde{\beta}),
  \]
  asymptotically.

Note

\[
0 = h(y, \tilde{\beta}) \\
\approx h(y, \beta) + \frac{\partial h(y, \beta)}{\partial \beta^t}(\tilde{\beta} - \beta) \\
= H^t(y - \mu) - H^tD(\tilde{\beta} - \beta).
\]
Thus,

\[ \tilde{\beta} - \beta \approx (H^tD)^{-1}H^t(y - \mu) \]

and

\[
\text{Var}(\tilde{\beta}) \approx (H^tD)^{-1}H^t\sigma^2VH(H^tD)^{-1t} = \sigma^2[D^tH(H^tVH)^{-1}H^tD]^{-1}.
\]

Note:

\[
\text{Var}(\hat{\beta}) \approx \sigma^2(D^tV^{-1}D)^{-1}.
\]

Now

\[
\text{Var}(\tilde{\beta}) - \text{Var}(\hat{\beta}) \geq 0
\]

\[
\iff [\text{Var}(\tilde{\beta})]^{-1} - [\text{Var}(\hat{\beta})]^{-1} \leq 0
\]

\[
\iff D^tV^{-1}D - D^tH(H^tVH)^{-1}H^tD \geq 0
\]

\[
\iff D^tV^{-1/2}[I - V^{1/2}H(H^tVH)^{-1}H^tV^{1/2}]V^{-1/2}D \geq 0
\]

\[
\iff I - V^{1/2}H(H^tVH)^{-1}H^tV^{1/2} \geq 0.
\]

- **Methods for constructing estimating equations**

  (i) **Unconditional construction**

  Suppose \( g_i(y, \theta), i = 1, \ldots, n, \) are \( n \) scaler estimating functions and \( \text{dim}(\theta) = p \). Let

\[
\begin{align*}
g & = (g_1, \ldots, g_n)^t, \\
V & = \text{Var}(g), \\
D & = -E\frac{\partial g}{\partial \theta^i}.
\end{align*}
\]
If \( \text{rank}(D) = p \), the estimating equations for estimating \( \theta \) is constructed as

\[
D^t V^{-1} g = 0.
\]

Remark: The idea of the construction is to combine \( n \) individual estimating functions together to form joint estimating equations. First, the individual estimating functions should be standardized to reflect their relative importance in the joint estimating equations. Second, the standardized individual functions can be considered as the derivatives of the log likelihood with respect to some implicit parameters which are functions of \( \theta \). The final estimating functions can be considered as the derivatives of the log likelihood with respect to \( \theta \).

**Example 1**: Quasi-likelihood estimating equations for dependent observations.

\[
y = (y_1, \ldots, y_n), \quad Ey_i = \mu_i, \quad \text{Var}(y) = \sigma^2 V(\mu),
\]

\[
\eta_i = g(\mu_i), \quad \eta_i = x_i^t \beta.
\]

Elementary estimating functions:

\[
g = y - \mu(\beta).
\]

Construction of estimating equations for \( \beta \):

\[
D = -E \frac{\partial g}{\partial \beta^t} = \frac{\partial \mu}{\partial \beta^t},
\]

\[
\frac{\partial \mu}{\partial \beta^t} V^{-1}(y - \mu(\beta)) = 0.
\]
(ii) **Conditional construction**

The method of un-conditional construction might fail because of 
(a) the singularity of $D$, or (b) the involvement of other nuisance parameters.

**Example 2:** First order autoregressive time series model.

$$Y_t = \theta Y_{t-1} + \epsilon_t, \quad Y_0 = 0,$$
$$\epsilon_t, \ i.i.d., \ \sim \ N(0, \sigma^2).$$

Elementary estimating functions:

$$g_t = Y_t - \theta Y_{t-1}.$$ 

But

$$D = -E \frac{\partial g}{\partial \theta} = E(y) = 0, \quad y = (Y_0, Y_1, \ldots, Y_{n-1})^t.$$ 

Method of conditional construction:

Let $g_i(y, \theta), i = 1, \ldots, n$, be $n$ elementary estimating functions. 
Let $A_i(y, \theta)$ be functions of $y$ and $\theta$. Suppose that 

$$E(g_i(y, \theta)|A_i(y, \theta)) = 0.$$ 

Let

$$V_i = \text{Var}(g_i(y, \theta)|A_i(y, \theta)),$$
$$D_{ir} = -E\left(\frac{\partial g_i(y, \theta)}{\partial \theta_r}\right)|A_i(y, \theta)),$$
$$V = \text{Diag}(V_1, \ldots, V_n), \quad D = (D_{ir})_{n \times p}.$$
If rank($D$) = $p$, the estimating equations for $\theta$ is constructed as

$$D^t V^{-1} g = 0, \quad g = (g_1, \ldots, g_n)^t.$$ 

**Example 2:** (cont.)

Let $A_t = Y_{t-1}$. Then

$$E(g_t|Y_{t-1}) = 0,$$
$$\text{Var}(g_t|Y_{t-1}) = \sigma^2,$$
$$D_t = -E\left(\frac{\partial g_t}{\partial \theta}|Y_{t-1}\right) = Y_{t-1}.$$ 

The estimating equation for $\theta$ is then

$$\sum_{t=1}^{n} Y_{t-1}(Y_t - \theta Y_{t-1}) = 0.$$ 

**Example 3:** Fieller-Creasy Problem:

Observations come in pairs: $(Y_{i1}, Y_{i2})$.

$Y_{i1}$ and $Y_{i2}$ are independent and normally distributed with the same variance $\sigma^2$, but

$$EY_{i1} = \mu_i, \quad EY_{i2} = \mu_i/\theta.$$ 

Parameter of interest: $\theta$.

Elementary estimating functions:

$$g_i(y_i, \theta) = Y_{i1} - \theta Y_{i2}.$$
Note:
\[ \frac{\partial g_i}{\partial \theta} = -Y_{i2}, \quad D_i = EY_{i2} = \mu_i/\theta_i. \]

The unconditional method fails here.

Try to find \( A_i(y_i, \theta) \) such that \( E(Y_{i2}|A_i(y_i, \theta)) \) is free of \( \mu_i \). Let \( A_i = aY_{i1} + bY_{i2}, a \) and \( b \) are free of \( \mu_i \).

\[
E(Y_{i2}|A_i(y_i, \theta)) = \mu_{Y_{i2}} + \frac{\text{Cov}(Y_{i2}, A_i)}{\text{Var}(A_i)}(A_i - \mu_{A_i})
\]

\[
= \frac{\mu_i}{\theta} + \frac{b\sigma^2}{(a^2 + b^2)\sigma^2}[A_i - (a\mu_i + b\mu_i/\theta)]
\]

\[
= \frac{\mu_i}{\theta} + \frac{b(a\theta + b)\mu_i}{a^2 + b^2 \theta} + \frac{b}{a^2 + b^2}A_i.
\]

Choose \( a \) and \( b \) such that
\[
\frac{b(a\theta + b)}{a^2 + b^2} = 1.
\]

This can be achieved by \( b = 1, \quad a = \theta \). Thus
\[
A_i = \theta Y_{i1} + Y_{i2}, \quad E(Y_{i2}|A_i) = \frac{A_i}{1 + \theta^2}.
\]

Note:
\[
\text{Cov}(g_i, A_i) = \text{Cov}(Y_{i1} - \theta Y_{i2}, \theta Y_{i1} + Y_{i2}) = 0.
\]

Hence
\[
E(g_i|A_i) = E(g_i) = 0, \quad \text{Var}(g_i|A_i) = \text{Var}(g_i) = (1 + \theta^2)\sigma^2.
\]
Estimating equation for $\theta$:

$$D^t V^{-1} g = \sum_{i=1}^{n} \frac{\theta Y_{i1} + Y_{i2}}{1 + \theta^2} \frac{1}{(1 + \theta^2)\sigma^2} (Y_{i1} - \theta Y_{i2})$$

$$= \sum_{i=1}^{n} \frac{(\theta Y_{i1} + Y_{i2})(Y_{i1} - \theta Y_{i2})}{(1 + \theta^2)^2 \sigma^2}$$

$$= 0.$$ 

Note: Let $u(\theta) = D^t V^{-1} g$. Then

$$u(\theta)|\{A_i : i = 1, \ldots, n\} \sim N(0, i(\theta)),$$

where

$$i(\theta) = \sum_{i=1}^{n} \frac{(Y_{i2} + \theta Y_{i1})^2}{(1 + \theta^2)^3 \sigma^2}.$$

A $100(1 - \alpha)\%$ confidence interval for $\theta$ can then be constructed as

$$\{\theta : \left| \frac{u(\theta)}{i^{1/2}(\theta)} \right| \leq z_{\alpha/2}\}.$$

**Example 4:** Estimation for megalithic stone rings.

Suppose $(Y_{i1}, Y_{i2}), i = 1, \ldots, n$, are observed Cartesian coordinates of points in a plane, assumed to lie on or near the circumference of a circle with center at $(\omega_1, \omega_2)$ and radius $\rho$. The center coordinates and the radius are to be estimated.
Assumed model:

\[ Y_{i1} = \omega_1 + R_i \cos \epsilon_i, \]
\[ Y_{i2} = \omega_2 + R_i \sin \epsilon_i, \]
\[ R_i > 0 \text{ i.i.d. } ER_i = \rho, \]
\[ \epsilon_i \text{ i.i.d. ind. } R_i. \]

Elementary estimating functions:

\[ g_i = \left[ (Y_{i1} - \omega_1)^2 + (Y_{i2} - \omega_2)^2 \right]^{1/2} - \rho \]
\[ = R_i(\omega_1, \omega_2) - \rho. \]

Elements of \( D \):

\[
\frac{\partial g_i}{\partial \theta} = \left\{ \begin{array}{l}
\frac{\partial g_i}{\partial \omega_1} = -\frac{Y_{i1}-\omega_1}{R_i}, \\
\frac{\partial g_i}{\partial \omega_2} = -\frac{Y_{i2}-\omega_2}{R_i}, \\
\frac{\partial g_i}{\partial \rho} = -1.
\end{array} \right.
\]

Let \( A_i = \epsilon_i \). Then

\[ E\left( \frac{\partial g_i}{\partial \theta} | A_i \right) = \frac{\partial g_i}{\partial \theta}. \]

Note: without conditioning on \( \epsilon_i \), \( E\left( \frac{\partial g_i}{\partial \theta} \right) \) are constant.

Estimating equation for \((\omega_1, \omega_2, \rho)\):

\[
\left\{ \begin{array}{l}
\sum_{i=1}^{n} \frac{Y_{i1}-\omega_1}{R_i} (R_i - \rho) = 0, \\
\sum_{i=1}^{n} \frac{Y_{i2}-\omega_2}{R_i} (R_i - \rho) = 0, \\
\sum_{i=1}^{n} (R_i - \rho) = 0.
\end{array} \right.
\]
Application to Avebury ring

Fig. 9.3 Diagram of the Avebury ring showing the fitted centres of arcs A, B and C, together with the joint fit and the fitted arc W. An approximate 99% confidence set is shown for the centre of arc A. Distances in feet.
Table 9.3  Stone number and position in the Avebury ring

<table>
<thead>
<tr>
<th>Arc C</th>
<th>Arc W</th>
<th>Arc A</th>
<th>Arc B</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>x</td>
<td>y</td>
<td>No.</td>
</tr>
<tr>
<td>1</td>
<td>733.7</td>
<td>44.0</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>659.7</td>
<td>28.0</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>624.2</td>
<td>19.3</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>588.4</td>
<td>13.9</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>551.6</td>
<td>12.3</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>515.1</td>
<td>9.5</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>478.0</td>
<td>16.6</td>
<td>15</td>
</tr>
<tr>
<td>16</td>
<td>243.5</td>
<td>183.0</td>
<td>37</td>
</tr>
<tr>
<td>17</td>
<td>216.3</td>
<td>205.0</td>
<td>38</td>
</tr>
<tr>
<td>18</td>
<td>188.9</td>
<td>229.8</td>
<td>39</td>
</tr>
<tr>
<td>19</td>
<td>163.5</td>
<td>255.5</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>140.0</td>
<td>285.0</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>120.6</td>
<td>305.7</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>103.1</td>
<td>323.1</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>85.9</td>
<td>344.0</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>61.8</td>
<td>371.3</td>
<td></td>
</tr>
</tbody>
</table>

\[
\hat{\omega}_1 = 530.8 \quad \hat{\omega}_1 = 1472.0 \quad \hat{\omega}_1 = 795.0 \quad \hat{\omega}_1 = 512.7 \\
\hat{\omega}_2 = 651.0 \quad \hat{\omega}_2 = 1553.4 \quad \hat{\omega}_2 = 516.5 \quad \hat{\omega}_2 = 533.1 \\
\hat{\rho} = 638.8 \quad \hat{\rho} = 1840.4 \quad \hat{\rho} = 782.8 \quad \hat{\rho} = 545.4 \\
\hat{\rho}^2 \hat{\sigma}^2 = 5.60 \quad \hat{\rho}^2 \hat{\sigma}^2 = 3.78 \quad \hat{\rho}^2 \hat{\sigma}^2 = 9.00 \quad \hat{\rho}^2 \hat{\sigma}^2 = 0.72
\]

†Data taken from Angell and Barber (1977).