Derivation of the Optimal LPQ Estimator in the Intercept-Only Model

1. Model Setup

Consider the intercept-only model

$$y_i = \mu + j_i, \quad i = 1, \dots, n,$$

with i.i.d. errors j_i satisfying

$$E[j_i] = 0, \quad \mathrm{Var}(j_i) = \sigma^2, \quad E[j_i^3] = m_3, \quad E[j_i^4] = m_4.$$

The usual unbiased estimator is $\bar{y} = \frac{1}{n}\sum_i y_i$ with $\mathrm{Var}(\bar{y}) = \sigma^2/n$. We seek an improved equivariant unbiased estimator of the form

$$\tilde{\mu} \;=\; \frac{1}{n}\sum_{i=1}^{n} y_i \;+\; B, \quad B = -\frac{1}{\sigma}\big(y^\top H y\big) + \sigma\,\mathrm{tr}(H),$$

where H is an $n \times n$ symmetric matrix to be chosen.

Equivariance/unbiasedness constraints: - Unbiasedness: $E[\tilde{\mu}] = \mu$. Since $E[y_i] = \mu$, the linear part $\frac{1}{n}\sum y_i$ is unbiased. For B,

$$E[B] = -\frac{1}{\sigma}E[y^\top H y] + \sigma\,\mathrm{tr}(H).$$

Write $y = \mu\mathbf{1} + j$. Then

$$y^\top H y = (\mu\mathbf{1} + j)^\top H(\mu\mathbf{1} + j) = \mu^2\,\mathbf{1}^\top H\mathbf{1} \;+\; 2\mu\,\mathbf{1}^\top H j \;+\; j^\top H j.$$

To ensure $E[B]$ does not depend on μ (so unbiased for all μ), require

$$H\,\mathbf{1} = 0 \quad \Longrightarrow \quad \mathbf{1}^\top H\mathbf{1} = 0,\; \mathbf{1}^\top H j = 0,$$

so $y^\top H y = j^\top H j$. Then

$$E[y^\top H y] = E[j^\top H j] = \sum_i H_{ii}E[j_i^2] + \sum_{i\neq j} H_{ij}E[j_i j_j] = \sigma^2\,\mathrm{tr}(H).$$

Hence $E[B] = -\frac{1}{\sigma}(\sigma^2\mathrm{tr}(H)) + \sigma\mathrm{tr}(H) = 0$. Thus unbiasedness holds whenever $H\mathbf{1} = 0$.

By symmetry of the i.i.d. setting, take H with

$$H_{ii} = a \quad (\forall i), \;\; H_{ij} = b \quad (\forall i \neq j), \;\; a + (n-1)b = 0 \;\Longrightarrow\; b = -\frac{a}{n-1}.$$

2. Variance decomposition

Write

$$\tilde{\mu} = \mu + A + B, \quad A = \frac{1}{n}\sum_{i=1}^{n} j_i, \quad B = -\frac{1}{\sigma}(j^\top H j) + \sigma\,\mathrm{tr}(H).$$

Since $E[A] = 0$, $E[B] = 0$,

$$\mathrm{Var}(\tilde{\mu}) = \mathrm{Var}(A) + \mathrm{Var}(B) + 2\,\mathrm{Cov}(A, B).$$

2.1. $\mathrm{Var}(A)$

$$\mathrm{Var}(A) = \mathrm{Var}\Big(\frac{1}{n}\sum_i j_i\Big) = \frac{\sigma^2}{n}.$$

2.2. $\mathrm{Cov}(A, B)$

$$\mathrm{Cov}\Big(A,\ -\frac{1}{\sigma}j^\top H j\Big) = -\frac{1}{n\sigma}\,\mathrm{Cov}\Big(\sum_{i=1}^{n} j_i,\ \sum_{k,\ell} H_{k\ell}\, j_k j_\ell\Big).$$

Only diagonal terms $k = \ell$ contribute under independence and zero mean:

$$\mathrm{Cov}\Big(\sum_i j_i,\ j^\top H j\Big) = \sum_{k=1}^{n} H_{kk} E[j_k^3] = (na)\, m_3.$$

Thus

$$\mathrm{Cov}(A, B) = -\frac{1}{n\sigma}\,(nam_3) = -\frac{a\, m_3}{\sigma}, \quad 2\,\mathrm{Cov}(A, B) = -\,\frac{2a\, m_3}{\sigma}.$$

2.3. $\mathrm{Var}(B)$

We need $\mathrm{Var}(j^\top H j)$. As before,

$$j^\top H j = a\sum_{i=1}^{n} j_i^2 \ +\ b\sum_{i\neq j} j_i j_j, \quad b = -\frac{a}{n-1}.$$

- $\mathrm{Var}\big(a\sum_i j_i^2\big) = a^2\sum_i \mathrm{Var}(j_i^2) = a^2 n\big(m_4 - \sigma^4\big)$. - $\mathrm{Var}\big(b\sum_{i\neq j} j_i j_j\big) = b^2\, n(n-1)\sigma^4$. No covariance between the two sums. Hence

$$\mathrm{Var}(j^\top H j) = a^2 n(m_4 - \sigma^4) + b^2 n(n-1)\sigma^4.$$

Thus

$$\mathrm{Var}(B) = \mathrm{Var}\Big(-\frac{1}{\sigma}j^\top H j\Big) = \frac{1}{\sigma^2}\,\mathrm{Var}(j^\top H j) = \frac{a^2 n(m_4 - \sigma^4) + b^2 n(n-1)\sigma^4}{\sigma^2}.$$

Substitute $b = -a/(n-1)$:

$$b^2 n(n-1)\sigma^4 = \frac{na^2\sigma^4}{n-1},$$

so

$$\mathrm{Var}(B) = \frac{a^2}{\sigma^2}\Big[n(m_4 - \sigma^4) + \frac{n\sigma^4}{n-1}\Big].$$

2.4. Total variance and minimization

Combine:

$$\mathrm{Var}(\tilde{\mu}) = \frac{\sigma^2}{n} \;+\; \frac{a^2}{\sigma^2}\Big[n(m_4-\sigma^4) + \tfrac{n\sigma^4}{n-1}\Big] \;-\; \frac{2a\,m_3}{\sigma}.$$

Let

$$C = n(m_4-\sigma^4) + \frac{n\sigma^4}{n-1}.$$

Then

$$\mathrm{Var}(\tilde{\mu}) = \frac{\sigma^2}{n} + \frac{a^2}{\sigma^2}C - \frac{2a\,m_3}{\sigma}.$$

Differentiate w.r.t. a:

$$\frac{d}{da}\Big(\frac{a^2}{\sigma^2}C - \frac{2a\,m_3}{\sigma}\Big) = 2\frac{a}{\sigma^2}C - 2\frac{m_3}{\sigma} = 0 \implies a = \frac{m_3\,\sigma}{C}.$$

Thus the optimal H has

$$H_{ii} = a = \frac{m_3\,\sigma}{C}, \quad H_{ij} = -\frac{a}{n-1} = -\frac{m_3\,\sigma}{(n-1)\,C},\ i \neq j.$$

Plugging back:

$$\frac{a^2}{\sigma^2}C = \frac{m_3^2\sigma^2}{C^2}\cdot\frac{C}{\sigma^2} = \frac{m_3^2}{C}, \quad -\frac{2am_3}{\sigma} = -\frac{2m_3^2}{C}.$$

Hence

$$\mathrm{Var}(\tilde{\mu}_{\mathrm{opt}}) = \frac{\sigma^2}{n} + \frac{m_3^2}{C} - \frac{2m_3^2}{C} = \frac{\sigma^2}{n} - \frac{m_3^2}{C}.$$

Recalling $C = n(m_4-\sigma^4) + \frac{n\sigma^4}{n-1}$, the final general formula is

$$\boxed{\mathrm{Var}(\tilde{\mu}_{\mathrm{opt}}) = \frac{\sigma^2}{n} \;-\; \frac{m_3^2}{n\big(m_4-\sigma^4\big) + \frac{n\sigma^4}{n-1}}.}$$

3. Remarks

- If $m_3 = 0$ (symmetric errors), then $a = 0$, so $H = 0$ and $\tilde{\mu} = \bar{y}$. - When $m_3 \neq 0$, one obtains variance reduction below σ^2/n, the amount governed by m_3^2 and m_4. - The matrix H remains exchangeable ($H_{ii} = a$, $H_{ij} = -a/(n-1)$) ensuring $H\mathbf{1} = 0$.

Asymptotic Normality of the LPQ Estimator in the Intercept-Only Model

1. Setup and form of the estimator

We observe $y_1, \ldots, y_n$ from the intercept-only model

$$y_i = \mu + j_i, \quad i = 1, \ldots, n,$$

where j_i are i.i.d. with

$$E[j_i] = 0, \quad \mathrm{Var}(j_i) = \sigma^2, \quad E[j_i^3] = m_3, \quad E[j_i^4] = m_4,$$

and assume $m_4 < \infty$. The usual estimator is $\bar{y} = \frac{1}{n} \sum_i y_i$. The LPQ estimator adds a quadratic adjustment:

$$\tilde{\mu} = \bar{y} \ + \ B, \quad B = -\frac{1}{\sigma}\big(y^\top H y\big) + \sigma \operatorname{tr}(H),$$

with H symmetric, $H\,\mathbf{1} = 0$, chosen to minimize variance. By symmetry one takes

$$H_{ii} = a, \quad H_{ij} = b \ (i \neq j), \quad a + (n-1)b = 0.$$

One shows (as in the finite-sample derivation) that

$$a = O\Big(\tfrac{1}{n}\Big), \quad b = -\frac{a}{n-1},$$

and that $y^\top H y = j^\top H j$. We now study asymptotic normality of $\tilde{\mu}$.

2. Rewriting B in terms of sample mean and sample variance

Define:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i, \quad \bar{j} = \frac{1}{n} \sum_{i=1}^{n} j_i = \bar{y} - \mu,$$

and the sample variance of the j_i:

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (j_i - \bar{j})^2.$$

We use that $H_{ii} = a$, $H_{ij} = b = -\frac{a}{n-1}$, and $H\mathbf{1} = 0$ implies

$$j^\top H j \ = \ a \sum_{i=1}^{n} j_i^2 \ + \ b \sum_{i \neq j} j_i j_j \ = \ a \sum_{i=1}^{n} j_i^2 \ - \ \frac{a}{n-1} \Big(\sum_{i \neq j} j_i j_j \Big).$$

Note

$$\sum_{i\neq j} j_i j_j = \Big(\sum_{i=1}^n j_i\Big)^2 - \sum_{i=1}^n j_i^2 = n^2\bar{j}^2 - \sum_{i=1}^n j_i^2.$$

Therefore

$$j^\top H j = a\sum_{i=1}^n j_i^2 - \frac{a}{n-1}\big(n^2\bar{j}^2 - \sum_{i=1}^n j_i^2\big) = a\Big[\sum_{i=1}^n j_i^2 + \frac{1}{n-1}\sum_{i=1}^n j_i^2 - \frac{n^2}{n-1}\bar{j}^2\Big] = a\Big[\frac{n}{n-1}\sum_{i=1}^n j_i^2 - \frac{n^2}{n-1}\bar{j}^2\Big].$$

But

$$\sum_{i=1}^n j_i^2 = \sum_{i=1}^n \big[(j_i-\bar{j})+\bar{j}\big]^2 = \sum_{i=1}^n (j_i-\bar{j})^2 \;+\; 2\bar{j}\sum_{i=1}^n (j_i-\bar{j}) \;+\; n\bar{j}^2 = nS^2 + n\bar{j}^2,$$

since $\sum (j_i-\bar{j}) = 0$. Hence

$$\sum_{i=1}^n j_i^2 = nS^2 + n\bar{j}^2.$$

Plug into $j^\top H j$:

$$j^\top H j = a\Big[\frac{n}{n-1}(nS^2+n\bar{j}^2) \;-\; \frac{n^2}{n-1}\bar{j}^2\Big] = a\frac{n}{n-1}\Big(nS^2 + n\bar{j}^2 - n\bar{j}^2\Big) = a\frac{n^2}{n-1}S^2.$$

Thus

$$j^\top H j = \frac{an^2}{n-1}\,S^2.$$

Next, $\mathrm{tr}(H) = na$. Therefore

$$B \;=\; -\frac{1}{\sigma}(j^\top H j) + \sigma\,\mathrm{tr}(H) = -\frac{1}{\sigma}\cdot\frac{an^2}{n-1}S^2 \;+\; \sigma\cdot(na).$$

Re-arrange:

$$B = -\frac{an^2}{\sigma(n-1)}\,S^2 \;+\; an\sigma.$$

We will later use the fact that $a = O(1/n)$, so $an \to 0$ and $an^2/(n-1) = O(1)$. More precisely, from the finite-sample optimal choice one shows $an \to 0$ but $an^2/(n-1) \to -\,m_3/(m_4-\sigma^4)$ as $n\to\infty$. For the asymptotic normality proof, we only need the order and the limit of $an^2/(n-1)$. Denote

$$\lambda_n \;=\; -\frac{an^2}{\sigma(n-1)}, \quad \rho_n \;=\; an\sigma.$$

Then

$$B = \lambda_n\,S^2 \;+\; \rho_n, \quad \tilde{\mu} - \mu = \bar{j} + B = \bar{j} + \lambda_n S^2 + \rho_n.$$

3. Asymptotic expansion

We assume the finite-sample optimal a satisfies, as $n \to \infty$,

$$a = -\frac{m_3\,\sigma^2}{n\,(m_4 - \sigma^4)} + o\Big(\tfrac{1}{n}\Big),$$

which implies

$$\lambda_n = -\frac{an^2}{\sigma(n-1)} = \frac{m_3\,\sigma^2}{n(m_4-\sigma^4)} \cdot \frac{n^2}{\sigma(n-1)} + o(1) = -\frac{m_3}{m_4-\sigma^4} + o(1),$$

and

$$\rho_n = an\sigma = O\Big(\tfrac{1}{n}\Big) \to 0.$$

Thus

$$\tilde{\mu} - \mu = \bar{j} + \lambda_n S^2 + \rho_n = \bar{j} + \lambda_n\big(S^2 - \sigma^2\big) + \lambda_n \sigma^2 + \rho_n.$$

But note $\lambda_n\sigma^2 + \rho_n = o(n^{-1/2})$, since $\lambda_n = O(1)$ and $\rho_n \to 0$. Also $S^2 - \sigma^2 = O_p(n^{-1/2})$ and $\bar{j} = O_p(n^{-1/2})$. Hence

$$\tilde{\mu} - \mu = \bar{j} + \lambda_n\big(S^2 - \sigma^2\big) + o_p(n^{-1/2}).$$

We set $\lambda = \lim_{n\to\infty} \lambda_n = -\frac{m_3}{m_4-\sigma^4}$. Then

$$\tilde{\mu} - \mu = \bar{j} + \lambda\big(S^2 - \sigma^2\big) + o_p(n^{-1/2}).$$

4. Joint CLT for $\bar{j}$ and S^2

Under i.i.d. j_i with $E[j_i] = 0$, $(j_i) = \sigma^2$, $E[j_i^4] < \infty$: - By the CLT, $\sqrt{n}\,\bar{j} \ \xrightarrow{d}\ N(0, \sigma^2)$. - Also $\sqrt{n}(S^2 - \sigma^2) \xrightarrow{d} N(0,\,(j_i^2))$, where $(j_i^2) = E[j_i^4] - \sigma^4 = m_4 - \sigma^4$. - And $(\bar{j},\ S^2 - \sigma^2) = \frac{1}{n}E[(j_i)(j_i^2 - \sigma^2)] = \frac{m_3}{n}$. Hence in the joint limit:

$$\sqrt{n}\begin{pmatrix} \bar{j} \\ S^2 - \sigma^2 \end{pmatrix} \ \xrightarrow{d}\ N\Big(0,\, \begin{pmatrix} \sigma^2 & m_3 \\ m_3 & m_4 - \sigma^4 \end{pmatrix}\Big).$$

5. Asymptotic normality of $\tilde{\mu}$

From the expansion

$$\tilde{\mu} - \mu = \bar{j} + \lambda\,(S^2 - \sigma^2) + o_p(n^{-1/2}),$$

we multiply by $\sqrt{n}$:

$$\sqrt{n}(\tilde{\mu} - \mu) = \sqrt{n}\,\bar{j} + \lambda\,\sqrt{n}(S^2 - \sigma^2) + o_p(1).$$

By the joint CLT above, the right-hand side converges in distribution to a normal:

$$\sqrt{n}(\tilde{\mu} - \mu) \ \xrightarrow{d}\ N(0,\, V),$$

where

$$V \;=\; \begin{pmatrix}1 & \lambda\end{pmatrix} \begin{pmatrix}\sigma^2 & m_3 \\ m_3 & m_4-\sigma^4\end{pmatrix} \begin{pmatrix}1 \\ \lambda\end{pmatrix} \;=\; \sigma^2 + 2\lambda m_3 + \lambda^2\,(m_4-\sigma^4).$$

Substitute $\lambda = -\frac{m_3}{m_4-\sigma^4}$:

$$V = \sigma^2 \;+\; 2\big(-\tfrac{m_3}{m_4-\sigma^4}\big)m_3 \;+\; \big(-\tfrac{m_3}{m_4-\sigma^4}\big)^2(m_4-\sigma^4).$$

Compute term by term: - $2\lambda m_3 = -2\frac{m_3^2}{m_4-\sigma^4}$. - $\lambda^2(m_4-\sigma^4) = \frac{m_3^2}{(m_4-\sigma^4)^2}\,(m_4-\sigma^4) = \frac{m_3^2}{m_4-\sigma^4}$. Hence

$$V = \sigma^2 \;-\; \frac{2m_3^2}{m_4-\sigma^4} \;+\; \frac{m_3^2}{m_4-\sigma^4} = \sigma^2 \;-\; \frac{m_3^2}{m_4-\sigma^4}.$$

Thus

$$\sqrt{n}(\tilde{\mu}-\mu) \;\xrightarrow{d}\; N\Big(0,\;\sigma^2 - \frac{m_3^2}{m_4-\sigma^4}\Big).$$

Equivalently,

$$\tilde{\mu} \;\approx\; N\Big(\mu,\;\tfrac{1}{n}\big[\sigma^2 - \tfrac{m_3^2}{m_4-\sigma^4}\big]\Big) \quad \text{for large } n.$$

This matches the leading term of the finite-sample variance reduction: $(\tilde{\mu}) = \frac{\sigma^2}{n} - \frac{m_3^2}{n(m_4-\sigma^4)+o(n)} \approx \frac{1}{n}(\sigma^2 - \frac{m_3^2}{m_4-\sigma^4})$.

6. Conclusion

Under finite fourth-moment assumptions, the LPQ estimator in the intercept-only model satisfies

$$\sqrt{n}(\tilde{\mu}-\mu) \;\xrightarrow{d}\; N\Big(0,\;\sigma^2 - \tfrac{m_3^2}{m_4-\sigma^4}\Big).$$

The key steps are:

- Rewrite the quadratic adjustment B in terms of the sample variance S^2 plus negligible terms.
- Use the joint CLT for $(\bar{j},\,S^2)$.
- Combine via a linear combination (since B is asymptotically linear in $S^2-\sigma^2$) to get a normal limit.

This completes the proof of asymptotic normality.