EFFICIENCY EVALUATION FOR QUASI-LINEAR INVARIANT PREDICTORS

by J. TIAGO de OLIVEIRA
Lisbon, Faculty of Sciences

1. Introduction.

Consider a sequence of random variables $X_1, \ldots, X_n, X_{n+1}, \ldots, X_{n+m}, \ldots$ from which we did obtain a sample $(X_1, \ldots, X_n)$.

In a previous paper (Tiago de Oliveira, 1966) we developed a method for obtaining quasi-linearly invariant predictors and quasi-linearly invariant prediction regions for a quasi-linearly invariant statistical function $Z = \psi(X_{n+1}, \ldots, X_{n+m})$ of the random variables, computed from the observed sample $(X_1, \ldots, X_n)$, when the model has only location and dispersion parameters. Recall that a statistical function $p(X_1, \ldots, X_n)$ is said to be quasi-linearly invariant if

$$p(\lambda + \delta X_1, \ldots, \lambda + \delta X_n) = \lambda + \delta p(X_1, \ldots, X_n)$$
$$(-\infty < \lambda < +\infty, \ 0 < \delta < +\infty)$$

The best (least-squares) quasi-linear predictor searched is a quasi-linear function p that minimizes

$$E^2 = \int_{-\infty}^{+\infty} \!\!\int [z - p(x_1, \ldots, x_n)]^2 \frac{1}{\delta^{n+1}}$$
$$\times \mathcal{L}\left(\frac{x_1 - \lambda}{\delta}, \ldots, \frac{x_n - \lambda}{\delta}; \frac{z - \lambda}{\delta}\right) dx_1 \ldots dx_n \, dz;$$

$$\frac{1}{\delta^{n+1}} \mathcal{L}\left(\frac{x_1 - \lambda}{\delta}, \ldots, \frac{x_n - \lambda}{\delta}; \frac{z - \lambda}{\delta}\right)$$

being the likehood of $(x_1, \ldots, x_n; z)$ because of the existence of a location (λ) and a dispersion (δ) parameter and of the quasi-linearity of ψ.

Research sponsored by the C. Gulbenkian Foundation.

As examples of quasi-linearly invariant statistics we can consider order statistics $X_{(i)}$ as the median, linear combinations of order statistics $\sum a_i X_{(i)}$ with $\sum a_i = 1$, as the average, etc.

When possible we will use, in the sequel, a condensed notation as $p(x)$, $\mathcal{L}(x; z)$, etc.

Using the quasi-linearity of p and passing to the reduced values (denoted also by x_i and z) we can write

$$E^2 = \delta^2 \int_{-\infty}^{\infty} \cdots \int [z - p(x)]^2 \mathcal{L}(x; z)\, dx\, dz = \delta^2 E_0^2$$

Remark that E and E_0 are equal quantities although measured in different units; E_0 is its value in standard ($\delta = 1$) units.

We have shown in (Tiago de Oliveira, 1966) that the best (least-squares) quasi-linear predictor is given by

$$p^*(x_1, \ldots, x_n)$$

$$= \frac{\iint_{\mathcal{D}} d\lambda\, d\delta\, \delta^{n+1} \int_{-\infty}^{+\infty} dz\, z\, \mathcal{L}(\lambda + \delta x_1, \ldots, \lambda + \delta x_n; \lambda + \delta z)}{\iint_{\mathcal{D}} d\lambda\, d\delta\, \delta^{n+1} \int_{-\infty}^{+\infty} dz\, \mathcal{L}(\lambda + \delta x_1, \ldots, \lambda + \delta x_n; \lambda + \delta z)}$$

where $\mathcal{D} = \{(\lambda, \delta) \mid -\infty < \lambda < +\infty; 0 < \delta < +\infty\}$.

Denoting by $\overline{\mathcal{L}}(x_1, \ldots, x_n)$ the marginal likehood of $(x_1, \ldots, x_n; z)$ and by

$$\mu(x_1, \ldots, x_n) = \int_{-\infty}^{+\infty} dz\, z\, \mathcal{L}(x_1, \ldots, x_n; z) / \overline{\mathcal{L}}(x_1, \ldots, x_n)$$

the conditional mean of z, we can write

$$p^*(x_1, \ldots, x_n)$$

$$= \frac{\iint_{\mathcal{D}} d\lambda\, d\delta\, \delta^{n-1} [\mu(\lambda + \delta x_1, \ldots, \lambda + \delta x_n) - \lambda] \times \overline{\mathcal{L}}(\lambda + \delta x_1, \ldots, \lambda + \delta x_n)}{\iint_{\mathcal{D}} d\lambda\, d\delta\, \delta^{n}\, \overline{\mathcal{L}}(\lambda + \delta x_1, \ldots, \lambda + \delta x_n)}$$

The fact that in many cases the best predictor is very difficult to compute necessitates naturally the use of different predictors. We will obtain now lower bounds for the mean square error of a predictor which will be useful for the definition and evaluation of efficiency. In fact, a computable decision technique with 70% or 80% efficiency is better than an uncomputable full-efficient technique.

2. **Predictor efficiency.**

In order to define efficiency of a (mean-square quasi-linear) predictor let us give a convenient form to the mean square error $E_0^2(p)$ of the predictor p.

As

$$
\begin{aligned}
E_0^2(p) &= \int_{-\infty}^{+\infty} \cdots \int [z - p(x)]^2 \, \mathcal{L}(x; z) \, dx \, dz \\
&= \int_{-\infty}^{+\infty} \cdots \int [z - \mu(x)]^2 \, \mathcal{L}(x; z) \, dx \, dz \\
&+ \int_{-\infty}^{+\infty} \cdots \int [\mu(x) - p(x)]^2 \, \overline{\mathcal{L}}(x) \, dx = \sigma^2 + D_0^2(p)
\end{aligned}
$$

where σ^2 denotes the variance of about the conditional mean $\mu(x_1, \ldots, x_n)$ and $D_0^2(p)$ the mean square distance between $\mu(x_1, \ldots, x_n)$ and the (proposed) predictor $p(x_1, \ldots, x_n)$. To obtain a lower bound for $E_0^2(p)$, as σ^2 is constant, it is sufficient to obtain a lower bound B_0^2 for $D_0^2(p)$.

Before computing B_0^2 let us consider two possible definitions of efficiency. If p^* denotes the best predictor as $E_0^2(p) \geqslant E_0^2(p^*)$ and, also $D_0^2(p) \geqslant D_0^2(p^*)$ efficiency can be defined as the quotient

$$
\frac{E_0^2(p^*)}{E_0^2(p)} \quad \text{or} \quad \frac{D_0^2(p^*)}{D_0^2(p)} .
$$

We will prefer the second definition on the following (heuristic) grounds: in the case of independence and $m = 1$ $(z = X_{n+1})$, σ^2 is a constant and if $D_0^2(p) \sim a/n^{\alpha}$ and $D_0^2(p^*) \sim a^*/n^{\alpha^*}$ we have, as $D_0^2(p) \geqslant D_0^2(p^*)$, $\alpha \leqslant \alpha^*$ so that $\dfrac{E_0^2(p^*)}{E_0^2(p)} \to 1$ and $D_0^2(p^*)/D_0^2(p) \sim a^*/a \cdot n^{\alpha-\alpha^*}$, the last one converging to zero if $\alpha < \alpha^*$, as it should be. Consequently we will take $D_0^2(p^*)/D_0^2(p)$ as the definition of efficiency for the predictor.

The results obtained from this definition can be easily translated to the other definition if adopted.

Similarly to what is done in estimation problem we will search lower bounds B_0^2 for $D_0^2(p)$, analogous to the Cramer-Rao or Kiefer bounds, see (Kendall and Stuart, II, 1961). A lower bound for the efficiency of p is then given by

$$B_0^2/D_0^2(p) .$$

Let us, now, obtain B_0^2. Denote

$$\phi(\lambda, \delta) = \int_{-\infty}^{+\infty} \cdots \int \mu(\lambda + \delta x_1, \ldots, \lambda + \delta x_n)\, \overline{\mathcal{L}}(x_1, \ldots, x_n)\, dx_1, \ldots, dx_n$$

We have

$$\int_{-\infty}^{+\infty} \cdots \int [p(x) - \mu(x)] \frac{1}{\delta^n} \overline{\mathcal{L}}\left(\frac{x-\lambda}{\delta}\right) dx$$

$$= \lambda + \delta \int_{-\infty}^{-\infty} \cdots \int p(x)\, \mathcal{L}(x)\, dx - \phi(\lambda, \delta)$$

on account of the quasi-linearity. Consequently we can write (to eliminate the expected value of p in the second member)

$$\int_{-\infty}^{+\infty} \cdots \int [p(x) - \mu(x)] \left[\frac{1}{\delta^n} \overline{\mathcal{L}}\left(\frac{x-\lambda}{\delta}\right) - \delta\, \overline{\mathcal{L}}(x)\right] dx$$

$$= \lambda - \phi(\lambda, \delta) + \delta\, \phi(0, 1)$$

The decomposition of the integrand as

$$[p(x) - \mu(x)] \sqrt{\overline{\mathcal{L}}(x)} \cdot \left[\frac{1}{\delta^n} \mathcal{L}\left(\frac{x-\lambda}{\delta}\right) - \delta\, \mathcal{L}(x)\right] / \sqrt{\overline{\mathcal{L}}(x)}$$

and the use of Schwarz's inequality gives

$$D_0^2(p) \geqslant \frac{[\lambda - \phi(\lambda, \delta) + \delta\, \phi(0, 1)]^2}{\displaystyle\int_{-\infty}^{+\infty} \cdots \int \frac{1}{\mathcal{L}(x)} \left[\frac{1}{\delta^n} \overline{\mathcal{L}}\left(\frac{x-\lambda}{\delta}\right) - \delta\, \overline{\mathcal{L}}(x)\right]^2 dx}$$

so that the lower bound is

$$B_0^2 = \sup_{(\lambda, \delta) \in \mathcal{D}} \frac{[\lambda - \phi(\lambda, \delta) + \delta\, \phi(0, 1)]^2}{\displaystyle\int_{-\infty}^{\infty} \cdots \int \frac{1}{\mathcal{L}(x)} \left\{\frac{1}{\delta^n} \overline{\mathcal{L}}\left(\frac{x-\lambda}{\delta}\right) - \delta\, \overline{\mathcal{L}}(x)\right\}^2 dx}$$

$$= \sup_{(\lambda,\,\delta)\,\epsilon\,\mathcal{D}} \frac{[\lambda - \phi(\lambda, \delta) + \delta\, \phi(0, 1)]^2}{W_n(\lambda, \delta) - 2\,\delta + \delta^2}$$

where

$$W_n(\lambda, \delta) = \int_{-\infty}^{+\infty} \cdots \int \frac{\overline{\mathcal{L}}^2\left(\dfrac{x-\lambda}{\delta}\right)}{\delta^{2n}\ \overline{\mathcal{L}}(x)}\, dx;$$

remark that for some values of (λ, δ) the denominator may be infinite so that the quotient may be zero, a fact that does not disturb the computation of sup.

The computation of this bound B_0^2, which imposes no regularity conditions, is in general a difficult one; under regularity conditions we can give a more usable result. If (λ_0, δ_0) is the point where the lower bound B_0^2 is attained, the function for which this is obtained is

$$\overline{p}(x_1, \ldots, x_n) = \mu(x_1, \ldots, x_n) - h\,\delta_0 + h\, \frac{\overline{\mathcal{L}}\left(\dfrac{x_1-\lambda_0}{\delta_0}, \ldots, \dfrac{x_n-\lambda_0}{\delta_0}\right)}{\delta_0^{\,n}\ \overline{\mathcal{L}}(x_1, \ldots, x_n)}$$

where h is a convenient constant, as it is well known from the equality condition in Schwarz's inequality. If $\overline{p}(x_1, \ldots, x_n)$ is quasi-linearly invariant then $\overline{p}$ is the best quasi-linear predictor searched.

As
$$\frac{[\lambda - \phi(\lambda, \delta) + \delta\, \phi(0, 1)]^2}{W_n(\lambda, \delta) - 2\,\delta + \delta^2}$$

presents an indeterminacy for $\lambda = 0$, $\delta = 1$ it seems natural to study its limit value when $\lambda \to 0$, $\delta \to 1$ in a special manner.

Taking $\delta = 1 + \beta\,\lambda$ and letting $\lambda \to 0$ $(\delta \to 1)$ we obtain

$$B'^2_0 = \sup_{\beta}\ \lim_{\substack{\lambda \to 0 \\ \delta = 1 + \beta\lambda}} \frac{[\lambda - \phi(\lambda, \delta) + \delta\, \phi(0, 1)]^2}{\displaystyle\int_{-\infty}^{+\infty} \cdots \int \frac{1}{\overline{\mathcal{L}}(x)} \left\{\frac{1}{\delta^n}\, \overline{\mathcal{L}}\left(\frac{x-\lambda}{\delta}\right) - \delta\, \overline{\mathcal{L}}(x)\right\}^2 dx}$$

the computation of which, in general, is not difficult; B'^2_0 as a (lower) bound is more manageable. Under regularity conditions we obtain

$$B'^2_0 = \sup_{\beta} \frac{[1 - \phi'_\lambda(0,1) - \beta\, \phi'_\delta(0,1) + \beta\phi(0,1)]^2}{\displaystyle\int_{-\infty}^{+\infty} \cdots \int \frac{1}{\overline{\mathcal{L}}(x)} \{\overline{A}(x) + \beta\, \overline{B}(x) + (n+1)\, \beta\, \overline{\mathcal{L}}(x)\}^2\, dx}$$

where

$$\overline{A}\,(x_1\,,\ \ldots,\ x_n) = \sum_1^n \frac{\partial\,\overline{\mathcal{L}}\,(x_1\,,\ \ldots,\ x_n)}{\partial\,x_i}$$

and

$$\overline{B}\,(x_1\,,\ \ldots,\ x_n) = \sum_1^n x_i\,\frac{\partial\,\overline{\mathcal{L}}\,(x_1\,,\ \ldots,\ x_n)}{\partial\,x_i}\,.$$

are supposed to exist and differentiation under the sign of integration is possible.

The bound B'^2_0, analogous to the Cramer-Rao bound and obtained under regularity conditions, also can be derived easily in a direct way, which we will sketch. The quasi-linearity relation for p gives

$$\int_{-\infty}^{+\infty} \cdots \int p\,(x)\,\overline{\mathcal{L}}\,(\frac{x-\lambda}{\delta})\,\frac{1}{\delta^n}\,dx = \lambda + \delta \int_{-\infty}^{+\infty} \cdots \int p\,(x)\,\overline{\mathcal{L}}\,(x)\,dx$$

and, deriving in order to λ and δ and rearranging we obtain:

$$\int_{-\infty}^{+\infty} \cdots \int \overline{A}\,(x)\,p\,(x)\,dx = -1$$

$$\int_{-\infty}^{+\infty} \cdots \int \overline{A}\,(x)\,dx = 0$$

$$\int_{-\infty}^{+\infty} \cdots \int [\overline{B}\,(x) + (n+1)\,\overline{\mathcal{L}}\,(x)]\,p\,(x)\,dx = 0$$

$$\int_{-\infty}^{+\infty} \cdots \int \overline{B}\,(x)\,dx = -x\,.$$

Multiplying the 3rd relation by β, summing to the 1st and using Schwarz's inequality we obtain:

$$D_0^2\,(p) \geqslant \sup_\beta \frac{\{1 + \int_{-\infty}^{+\infty} \cdots \int \mu\,(x)\,[\overline{A}\,(x) + \beta\,\overline{B}\,(x) + (n+1)\,\beta\,\overline{\mathcal{L}}\,(x)]\,dx\}^2}{\int_{-\infty}^{+\infty} \cdots \int \frac{1}{\overline{\mathcal{L}}\,(x)}\,[\overline{A}\,(x) + \beta\,\overline{B}\,(x) + (n+1)\,\beta\,\overline{\mathcal{L}}\,(x)]^2\,dx}$$

which is B'^2_0, because

$$\phi'_\lambda\,(0,1) = -\int_{\infty}^{+\infty}\cdots\int \overline{A}\,(x)\;\mu\,(x)\;dx$$

$$\phi'_\delta\,(0,1) = -\int_{-\infty}^{+\infty}\cdots\int [n\;\overline{\mathcal{L}}\,(x) + \overline{B}\,(x)]\;\mu\,(x)\;dx$$

$$\phi\,(0,\;1) = \int_{\infty}^{\infty}\cdots\int \mu\,(x)\;\overline{\mathcal{L}}\,(x)\;dx\;.$$

Using the conditions obtained on $\overline{A}$ and $\overline{B}$ and denoting by

$$a = 1 + \int_{-\infty}^{+\infty}\cdots\int \mu\,(x)\;\overline{A}\,(x)\;dx$$

$$b = \int_{-\infty}^{+\infty}\cdots\int \mu\,(x)\;[\overline{B}\,(x) + (n+1)\;\overline{\mathcal{L}}\,(x)]\;dx$$

$$p = \int_{-\infty}^{+\infty}\cdots\int \frac{\overline{A}^2\,(x)}{\overline{\mathcal{L}}\,(x)}\;dx$$

$$q = \int_{-\infty}^{+\infty}\cdots\int \frac{\overline{A}\,(x)\;\overline{B}\,(x)}{\overline{\mathcal{L}}\,(x)}\;dx$$

$$r = \int_{-\infty}^{+\infty}\cdots\int \frac{\overline{B}^2\,(x)}{\overline{\mathcal{L}}\,(x)}\;dx - (n^2 - 1)$$

we have

$$B'^2_0 = \sup_{\beta}\;\frac{(a + b\,\beta)^2}{p + 2\,q\,\beta + r\,\beta^2} = \frac{a^2\,r + b^2\,p + 2\,abq}{pr - q^2}$$

this bound being obtained for

$$\overline{p}\,(x) = \mu\,(x) + \frac{bq - ar}{pr - q^2}\;\frac{\overline{A}\,(x)}{\overline{\mathcal{L}}\,(x)} + \frac{aq - bp}{pr - q^2}\;(\frac{\overline{B}\,(x)}{\overline{\mathcal{L}}\,(x)} + n + 1)$$

as it is well known from the conditions of equality for Schwarz's inequality. The function $\overline{p}\,(x)$ is the solution if quasi-linearly invariant.

In the case of independence we can, in general, give simple formulas as usual. In that case we have

$$\mathcal{L}\,(x;\;z) = \overline{\mathcal{L}}\,(x)\;g_m\,(z) \qquad \text{and} \qquad \overline{\mathcal{L}}\,(x) = f\,(x_1)\;\cdots\;f\,(x_n).$$

Consequently we have

$$\mu(x) = \mu_m \text{ (const)}$$

so that

$$\phi(\lambda, \delta) = \mu_m$$

We obtain, putting

$$W(\lambda, \delta) = \int_{-\infty}^{+\infty} \frac{1}{\delta^2} f^2 \left(\frac{x-\lambda}{\delta}\right) / f(x)\, dx,$$

$$B'^2_0 = \sup_{(\lambda, \delta) \in \mathcal{D}} \frac{(\lambda + \delta\mu_m - \mu_m)^2}{W^n(\lambda, \delta) - 2\delta + \delta^2}$$

B'^2_0 is easily obtained because

$$a = 1, \; b = \mu_m$$

$$p = n \int_{-\infty}^{+\infty} \frac{f'(x)^2}{f(x)}\, dx = np_1,$$

$$q = n \int_{-\infty}^{+\infty} x \frac{f'(x)^2}{f(x)}\, dx = nq_1$$

$$r - 1 = n \int_{-\infty}^{+\infty} x^2 \frac{f'(x)^2}{f(x)}\, dx = n(r_1 - 1)$$

if

$$\int_{-\infty}^{+\infty} f'(x)\, dx = 0 \text{ and } \int_{-\infty}^{+\infty} x f'(x)\, dx = 1 \text{ (regularity conditions).}$$

Using the expressions of p, q and r in p_1, q_1 and r_1 we obtain

$$B'^2_0 = \frac{1 + n(r_1 - 1 + p_1\mu_m^2 - 2\mu_m q_1)}{n[p_1 + n(p_1 r_1 - q_1^2 - p_1)]}$$

$$\sim \frac{p_1\mu_m^2 - 2\mu_m q_1 + r_1 - 1}{n(p_1 r_1 - q_1^2 - p_1)}$$

which shows that B'^2_0 is of order n^{-1}, for large samples, as we could expect.

In the case of normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

and one step ($m = 1$) prediction we have

$$b = \mu_m = 0, \quad p_1 = 1, \quad q_1 = 0, \quad r_1 = 3$$

so that

$$B'^2_0 = 1/n .$$

In the case of Gumbel distribution $f(x) = e^{-x} \exp(-e^{-x})$ and one-step prediction we obtain

$$b = \mu_m = \gamma, \quad p_1 = 1, \quad q_1 = \gamma - 1, \quad r_1 = \frac{\pi^2}{6} + (1 - \gamma)^2$$

so that

$$B'^2_0 = \frac{1 + n(1 + \pi^2/6)}{n\,[1 + n(\pi^2/6)]} \sim \frac{1}{n}\left(1 + \frac{6}{\pi^2}\right) \cong \frac{1.61}{n}$$

In the case of exponential, uniform, Weibull distributions, B'^2_0 can not be computed and it is, in many cases, possible to compute B^2_0.

3. **Final remarks.**

The fact that, for the independence case, the lower bound B'^2_0 is of order n^{-1} suggests the use of moments for prediction problem, that is, to take $p(x) = \bar{x} + \theta s$ as predictor. It is very easy to obtain in that case the best values for θ or, at least, the best asymptotic values. General formulas can be developed but it seems better to obtain them for each case. Expressions for a lower bound of the length of a prediction region can be deduced in an analogous way. This bound being proportional to ω^2 (ω, the prediction level) is not very sharp and useful, as it is easily seen for the normal case.

We profit this paper to make a correction relating to our previous paper. The proof of the prediction region is not correct although the result is. The correct proof is: we decompose $W(\varphi) = \omega = \omega' + \omega''$ ($\omega', \omega'' \geqslant 0$) according to $x_1 < x_2$ and $x_1 > x_2$; the average length is similarly decomposed and the minimization procedure in each of the half-spaces (Neyman-Pearson lemma) leads to subregions defined by constants k' and k''. Its equality is proved comparing with the constants for the decomposition $x_1 < x_3$ and $x_1 > x_3$, for instance.

REFERENCES

Kendall, M.G. and Stuart, A. (1961) — The Advanced Theory of Statistics, vol. II, C. Griffin & Co.

Tiago de Oliveira, J. (1966) — Quasi-linearly invariant prediction, *Ann. Math. Statist.*, vol. 37, p. 6.