ESAIM: Probability and Statistics December 2000, Vol. 4, 229–231
URL: `http://www.emath.fr/ps/`

A LEMMA ON PROXIMITY OF VARIANCES AND EXPECTATIONS *

Doron Sonsino[1]

Abstract. We define a notion of delta-variance maximization and show it implies epsilon-proximity in expactations.

AMS Subject Classification. 60A99.

Received March 19, 2000.

We present a lemma stipulating that when the variance of each element in a collection of random variables is maximal with respect to some larger class of random variables, then the corresponding expectations must be very close.

First, we formally define our notion of a large class of random variables.

Definition. A collection $\mathcal{G}$ of random variables on a probability space $(\Omega, \mathcal{F}, P)$ is *CPC (Closed under Piecewise Compositions)* if for every $y, y' \in \mathcal{G}$ and every measurable $E \in \mathcal{F}$, $\quad y \cdot I_E + y' \cdot I_{E^C} \in \mathcal{G}$[2].

In words, the collection $\mathcal{G}$ is closed under piecewise compositions if whenever y and y' are in $\mathcal{G}$ and E is a measurable set, the random variable that takes the value $y(\omega)$ when ω belongs to E and takes the value $y'(\omega)$ elsewhere, is in the collection $\mathcal{G}$ as well.

We say that the collection of random variables $\mathcal{G}$ is a *CPC–extension* of the collection $\mathcal{D}$ if $\mathcal{G}$ is a CPC–collection of random variables that contains $\mathcal{D}$ (where all the random variables in both collections are defined on the same probability space $(\Omega, \mathcal{F}, \mathcal{P})$).

We now define the notion of δ–variance maximization.

Definition. A collection $\mathcal{D}$ of random variables in $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ is *δ–variance maximizing* if there is a CPC–extension of $\mathcal{D}$, $\mathcal{G}$, and a finite k such that:

(1) $\quad \mathbf{Var}(y) \leq k \quad \forall y \in \mathcal{G}$, and
(2) $\quad \mathbf{Var}(y) \geq k - \delta \quad \forall y \in \mathcal{D}$.

Put differently, the collection of random variables $\mathcal{D}$ is δ–variance maximizing if there is a large (in the sense of CPC) collection of random variables $\mathcal{G}$ that contains $\mathcal{D}$ such that the variance of each of the random variables in $\mathcal{D}$ is within δ from the supremum of variances over the larger collection $\mathcal{G}$.

Note that if y and y' belong to a δ–variance maximizing collection of random variables, then $|\mathbf{Var}(y) - \mathbf{Var}(y')| \leq \delta$ and the two random variables have similar variances. Clearly, closeness in variances alone is not

Keywords and phrases: Variance, expectation.

* *I thank Dov Monderer, Haim Raizman and Ishy Weissman for comments and discussions.*

[1] Technion, Israel Institute of Technology, Haifa, Israel; e-mail: `sonsino@ie.technion.ac.il`

[2] Measurability of $y \cdot I_E + y' \cdot I_{E^C}$ follows directly from standard arguments.

strong enough to induce similarity in expectations. The MVSE lemma[3] however says that the stronger concept of δ–variance maximization is sufficient for this purpose.

The MVSE lemma. Let $\mathcal{D}$ be a collection of random variables on a non atomic probability space $(\Omega, \mathcal{F}, P)$. If $\mathcal{D}$ is δ–variance maximizing then $|\mathbf{E}[y] - \mathbf{E}[y']| \le 2\sqrt{\delta}$ for every $y, y' \in \mathcal{D}$.

Proof. Let $\mathcal{D}$ be a δ–variance maximizing collection of random variables on a non atomic probability space $(\Omega, \mathcal{F}, P)$. Fix y, y' in $\mathcal{D}$ and assume w.l.g. that $\mathbf{E}\,[y] > \mathbf{E}\,[y']$. Assume by way of contradiction that $\mathbf{E}\,[y] - \mathbf{E}\,[y'] = 2\sqrt{\epsilon} > 2\sqrt{\delta}$.

Let

$$A = \{\omega \in \Omega \ \mid \ \big(y(\omega) - \mathbf{E}\,[y] + \sqrt{\epsilon}\big)^2 \ge \big(y'(\omega) - \mathbf{E}\,[y'] - \sqrt{\epsilon}\big)^2\}$$
$$B = A^C = \{\omega \in \Omega \ \mid \ \big(y(\omega) - \mathbf{E}[y] + \sqrt{\epsilon}\big)^2 < \big(y'(\omega) - \mathbf{E}[y'] - \sqrt{\epsilon}\big)^2\}.$$

Observe that (by standard arguments) A and B are measurable w.r.t $\mathcal{F}$.

Note that since $\mathbf{E}\,[y] - \mathbf{E}\,[y'] = \int (y - y')\, dP = 2\sqrt{\epsilon}$, it is either the case that

$$\int_A (y - y')\, dP \ge \sqrt{\epsilon}, \quad \text{or} \tag{1}$$
$$\int_B (y - y')\, dP \ge \sqrt{\epsilon}. \tag{2}$$

In case 1, let E be a measurable subset of A such that

$$\int_E (y - y')\, dP = \sqrt{\epsilon}. \tag{3}$$

The existence of such a measurable E follows from the assumption that P is non atomic (see Billingsley [1], 2.17, p. 31).

Set $z = y \cdot I_E + y' \cdot I_{E^C}$ and observe that z must belong to any CPC–extension of $\mathcal{D}$. Thus, the assumptions that $\mathcal{D}$ is δ–variance maximizing implies that

$$\mathbf{Var}\,(z) \le \mathbf{Var}\,(y) + \delta \quad \text{and} \quad \mathbf{Var}\,(z) \le \mathbf{Var}\,(y') + \delta. \tag{4}$$

But note that by definitions of z and E,

$$\mathbf{E}\,[z] = \int z\, dP = \int_E y\, dP + \int_{E^C} y'\, dP = \int_\Omega y'\, dP + \int_E (y - y')\, dP = \mathbf{E}\,[y'] + \sqrt{\epsilon},$$

and similarly

$$\mathbf{E}\,[z] = \int_\Omega y\, dP - \int_{E^C} (y - y')\, dP = \int_\Omega y\, dP - \int_\Omega (y - y')\, dP + \int_E (y - y')\, dP = \mathbf{E}\,[y] - \sqrt{\epsilon}.$$

Thus,

$$\mathbf{Var}\,[z] = \int \big(z - E[z]\big)^2\, dP = \int_E (\,y - \mathbf{E}\,[y] + \sqrt{\epsilon}\,)^2\, dP + \int_{E^C} (\,y' - \mathbf{E}\,[y'] - \sqrt{\epsilon}\,)^2\, dP,$$

[3] MVSE stands for Maximal Variance Similar Expectations.

and since $(y - \mathbf{E}\,[y] + \sqrt{\epsilon}\,)^2 \geq (y' - \mathbf{E}\,[y'] - \sqrt{\epsilon}\,)^2$ on $E \subseteq A$,

$$\mathbf{Var}\,[z] \geq \int_\Omega (y' - \mathbf{E}\,[y'] - \sqrt{\epsilon}\,)^2 \, dP = \int_\Omega (y' - \mathbf{E}\,[y']\,)^2 \, dP + \epsilon = \mathbf{Var}\,[y'] + \epsilon > \mathbf{Var}\,[y'] + \delta,$$

which contradicts (4) and proves that case (1) is impossible.

In a very similar way we may argue that case (2) leads to a contradiction as well: assume by way of contradiction that the condition in (2) holds. Let E be a measurable subset of B such that

$$\int_E (y - y')\, dP = \sqrt{\epsilon}. \tag{5}$$

(Existence follows again from the assumption that P is non atomic.)

Let $z = y' \cdot I_E + y \cdot I_{E^C}$.

Note (as above) that

$$\mathbf{E}\,[z] = \mathbf{E}\,[y] - \sqrt{\epsilon} = \mathbf{E}\,[y'] + \sqrt{\epsilon}.$$

Thus,

$$\begin{aligned}\mathbf{Var}[z] &= \int \big(z - E[z]\big)^2 dP = \int_E (y' - E[y'] - \sqrt{\epsilon})^2 dP + \int_{E^C} (y - \mathbf{E}[y] + \sqrt{\epsilon})^2 dP \\ &\geq \int_\Omega (y - \mathbf{E}[y] + \sqrt{\epsilon})^2 dP = \int_\Omega \big(y - \mathbf{E}[y]\big)^2 dP + \epsilon = \mathbf{Var}[y] + \epsilon > \mathbf{Var}[y] + \delta,\end{aligned}$$

which again contradicts (4) and proves that case (2) is impossible as well. □

Reference

[1] P. Billingsley, *Probability and Measure.* 2nd edition. John Wiley & Sons, New York (1986).