Statistics & Probability Letters 15 (1992) 41–46
North-Holland
3 September 1992

# On the stability problem for conditional expectation

Wlodzimierz Bryc
*Department of Mathematics, University of Cincinnati, OH, USA*

Wlodzimierz Smolenski *
*Department of Mathematics, The University of Kansas, Lawrence, KS, USA*

Received April 1991
Revised February 1992

*Abstract*: The behavior of the conditional expectation $E\{X \mid Y\}$ under a small perturbation $Z$ of the conditioning random variable $Y$ is analyzed. We show that if $Y$ and $Z$ are independent then $E\{X \mid Y+\varepsilon Z\}$ converges to $E\{X \mid Y\}$ in mean as $\varepsilon \to 0$ for all integrable $X$, provided the distribution of $Y$ is absolutely continuous. We also show that the limit is $E\{X \mid Y, Z\}$ rather than $E\{X \mid Y\}$, i.e., there is no stability, when $Y$ is a discrete (i.e., countably valued) random variable. Finally, we show that in general $E\{X \mid Y+\varepsilon Z\}$ might have no limit in distribution as $\varepsilon \to 0$.

*AMS 1980 Subject Classification*: 60H99.

*Keywords*: Conditional expectations; stability; perturbation.

## 1. Introduction

In this note we investigate the stability of the conditional expectation $E\{X \mid Y\}$ under a small additive perturbation by a random variable $Z$, i.e., we study the behavior of $E\{X \mid Y+\varepsilon Z\}$ as $\varepsilon \to 0$. In the case when $Y$ and $Z$ are independent real valued and the convergence is understood in probability it is possible to answer the stability problem completely. Namely, we show that $E\{X \mid Y+\varepsilon Z\}$ converges to $E\{X \mid Y\}$ as $\varepsilon \to 0$, if $Y$ is absolutely continuous, $E\{X \mid Y+\varepsilon Z\}$ converges to $E\{X \mid Y, Z\}$ as $\varepsilon \to 0$, if $Y$ is discrete, and there is no limit in general. The same question can be asked also when $Y$ and $Z$ are $\mathbb{R}^d$-valued random variables or when $Y$ and $Z$ are replaced by stochastic processes $\{Y_t\}$; however, the information that we have in those cases is far less complete and is left out of this note.

The following result gives a sufficient condition for the stability of the conditional expectation under independent perturbations.

**Theorem 1.1.** *If $Y$, $Z$ are independent, the distribution of $Y$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and $E\{|X|\} < \infty$, then*

$$E\{X \mid Y+\varepsilon Z\} \to E\{X \mid Y\} \quad \text{in } L_1\{\mathrm{d}P\} \text{ as } \varepsilon \to 0. \tag{1.1}$$

*Correspondence to*: Wlodzimierz Bryc, Department of Mathematics, University of Cincinnati, Cincinnati, OH 45221-0025, USA.
* Supported in part by GRF grant from the University of Kansas.

**Remark 1.** Suitable modifications in the proof of Theorem 1 show that the conclusion holds also for dependent random variables $Y$, $Z$, provided they have joint density.

**Remark 2.** In general, to show that (1.1) holds for all integrable $X$, it is enough to consider all $X$ of the form $X=\phi(Z)$, where $\phi:\mathbb{R}\to\mathbb{R}$ is bounded and uniformly continuous. Indeed, by the standard approximation argument, the finite sums $\Sigma\phi_i(Z)\Psi_i(Y)$, where $\phi_i$, $\Psi_i$ are bounded and uniformly continuous, are dense in $L_1(\sigma(Y, Z), \mathrm{d}P)$. For a single term of the sum we have

$$E\{\phi(Z)\Psi(Y)\mid Y+\varepsilon Z\}=\Psi(Y+\varepsilon Z)E\{\phi(Z)\mid Y+\varepsilon Z\}+\theta_\varepsilon,$$

where

$$\theta_\varepsilon=E\{\phi(Z)(\Psi(Y)-\Psi(Y+\varepsilon Z))\mid Y+\varepsilon Z\}\to 0 \quad \text{in } L_1(\mathrm{d}P) \text{ as } \varepsilon\to 0.$$

Therefore $E\{\phi(Z)\Psi(Y)\mid Y+\varepsilon Z\}\to\Psi(Y)E\{\phi(Z)\mid Y\}$, provided (1.1) is established for $X=\phi(Z)$.

**Remark 3.** The argument given above also shows that (1.1) holds for all $\sigma(Y)$-measurable integrable $X$. As a consequence, one easily gets that if $Y$ has a density, $E\{|Z|^p\}<\infty$ for some $p\geqslant 1$ and $E\{Z\}=0$, then $E\{Y\mid Y+\varepsilon Z\}=Y+\varepsilon Z+\mathrm{o}(\varepsilon)$, where $\|\mathrm{o}(\varepsilon)\|_p/|\varepsilon|\to 0$ as $\varepsilon\to 0$. Indeed,

$$E\{Y\mid Y+\varepsilon Z\}=Y+\varepsilon Z-\varepsilon E\{Z\mid Y+\varepsilon Z\} \quad \text{and} \quad E\{Z\mid Y+\varepsilon Z\}\to E\{Z\}$$

by Theorem 1.1. Szablowski (1986, Corollary 1.2) shows that if one conditions $Z$ rather than $Y$, then the representation

$$E\{Z\mid Y+\varepsilon Z\}=E\{Z\}+\varepsilon\beta(Y+\varepsilon Z)+\mathrm{o}(\varepsilon)$$

permits one to determine the distribution of $Y$, which then has to be absolutely continuous.

The following result shows that there is no stability in the discrete case. No assumption of independence is needed.

**Theorem 1.2.** *If $Y$ is discrete and $E\{|X|\}<\infty$ then*

$$E\{X\mid Y+\varepsilon Z\}\to E\{X\mid Y, Z\} \quad \text{in } L_1(\mathrm{d}P) \text{ as } \varepsilon\to 0 \text{ for all } Z. \tag{1.2}$$

**Remark 4.** If in addition either
(i) $Z$ is discrete, or
(ii) $Z$ is bounded and $Y$ takes a finite number of values,
then the convergence in (1.2) is also in the most sure sense.

The following observation complements Theorems 1.1 and 1.2.

**Proposition 1.3.** *Suppose $Y$, $Z$ are independent and identically distributed. If (1.1) holds for all integrable $X$, then either $Y$ is continuous (i.e., atomless), or $Y$ is deterministic.*

Proposition 1.3 gives a necessary condition for (1.1). The following example shows that there are i.i.d. continuous (but not absolutely continuous) $Y$, $Z$ such that $E\{X\mid Y+\varepsilon Z\}$ has no limit in distribution as $\varepsilon\to 0$.

**Example.** Let $\xi_k$, $\eta_k$ be independent $\{0, 1\}$-valued random variables. Suppose

$$P(\xi_{2k}=1)=P(\eta_{2k}=1)=p_1 \quad \text{and} \quad P(\xi_{2k+1}=1)=P(\eta_{2k+1}=1)=p_2 \qquad \text{for all } k,$$

where $0 < p_1 < p_2 < 1$ are fixed. Let

$$X = \xi_1, \qquad Y = \sum_{k=1}^{\infty} \eta_k / 3^k, \qquad Z = \sum_{k=1}^{\infty} \xi_i / 3^k.$$

Clearly, $Y$ and $Z$ are independent and identically distributed. Moreover, from the uniqueness of the ternary expansion in this case we see that

$$\sigma\left\{Y + \frac{1}{3^n} Z\right\} = \sigma\{\eta_1, \eta_2, \ldots, \eta_n, \eta_{n+1} + \xi_1, \eta_{n+2} + \xi_2, \ldots\}.$$

Therefore

$$E\left\{X \mid Y + \frac{1}{3^n} Z\right\} = E\{\xi_1 \mid \eta_{n+1} + \xi_1\} = \begin{cases} 0 & \text{if } \eta_{n+1} + \xi_1 = 0, \\ \dfrac{p_1(1-q)}{p_1 + q - 2p_1 q} & \text{if } \eta_{n+1} + \xi_1 = 1, \\ 1 & \text{if } \eta_{n+1} + \xi_1 = 2, \end{cases}$$

where $q = p_1$ if $n$ is even and $q = p_2$ if $n$ is odd. In particular, $E\{X \mid Y + (1/3^n)Z\}$ has no limit in distribution as $n \to \infty$.

## 2. Proofs

**Proof of Theorem 1.1.** As it was pointed out in Remark 2, it is enough to establish (1.1) for $X = f(Z)$, where $f(\cdot)$ is a bounded continuous function. Furthermore, since the family of conditional expectations of a fixed random variable $X$ with respect to different $\sigma$-fields is uniformly integrable, we need only to show convergence in probability.

Let $g(y)$ be the density of $Y$. By $F_X$, $F_Y$, $F_Z$ we denote the cumulative distribution functions of $X$, $Y$, $Z$; $\lambda(\cdot)$ denotes the Lebesgue measure and $\mu \otimes F$ is the product of measures $\mu$ and $\mathrm{d}F$. Clearly,

$$E\{X \mid Y\} = \int f(z) F_Z(\mathrm{d}z)$$

and

$$E\{X \mid Y + \varepsilon Z\} = \frac{\int f(z) g(Y + \varepsilon Z - \varepsilon z) F_Z(\mathrm{d}z)}{\int g(Y + \varepsilon Z - \varepsilon z) F_Z(\mathrm{d}z)}.$$

Since $f$ is bounded, we have

$$\left| \int f(z) g(Y + \varepsilon Z - \varepsilon z) F_Z(\mathrm{d}z) - \int f(z) g(Y) F_Z(\mathrm{d}z) \right|$$
$$\leqslant \| f \|_\infty \int | g(Y + \varepsilon Z - \varepsilon z) - g(Y) | F_Z(\mathrm{d}z).$$

Therefore, to show that $E\{X \mid Y + \varepsilon Z\} \to E\{X \mid Y\}$ in probability, it is enough to show that

$$\int | g(Y + \varepsilon Z - \varepsilon z) - g(Y) | F_Z(\mathrm{d}z) \to 0 \quad \text{in probability.} \tag{2.1}$$

To this end, for each $\delta > 0$ let $h(\cdot) = h_\delta(\cdot)$ be a uniformly continuous probability density function such that

$$\int |g(y) - h(y)| \,\mathrm{d}y < \delta^2.$$

For real $\varepsilon$ let

$$A_\varepsilon = \left\{ (y, z) \colon \int |g(y + \varepsilon z - \varepsilon u) - h_\delta(y + \varepsilon z - \varepsilon u)| F_Z(\mathrm{d}u) > \delta \right\}.$$

By Chebyshev's inequality,

$$(\lambda \otimes F_Z)(A_\varepsilon) \leqslant \frac{1}{\delta} \iiint |g(y + \varepsilon z - \varepsilon u) - h(y + \varepsilon z - \varepsilon u)| \,\mathrm{d}y F_Z(\mathrm{d}u) F_Z(\mathrm{d}z) < \delta.$$

Let $\rho > 0$ be arbitrary. Since $F_Y \otimes F_Z$ is absolutely continuous with respect to $\lambda \otimes F_Z$, therefore

$$P((Y, Z) \in A_\varepsilon) < \rho \tag{2.2}$$

for all $\varepsilon$ and for all $\delta$ small enough, say for all $\delta < \delta_0(\rho)$. Let

$$B = B(\varepsilon, \delta) = \left\{ (y, z) \colon \int |g(y + \varepsilon z - \varepsilon u) - g(y)| F_Z(\mathrm{d}u) > 4\delta \right\},$$

$$C = C(\varepsilon, \delta) = \left\{ (y, z) \colon \int |h(y + \varepsilon z - \varepsilon u) - h(y)| F_Z(\mathrm{d}u) > 2\delta \right\}.$$

To prove (2.1) it is enough to show that $P((Y, Z) \in B) < 3\rho$ as $\varepsilon \to 0$. Clearly

$$P((Y, Z) \in B) \leqslant P((Y, Z) \in A_0) + P((Y, Z) \in A_\varepsilon) + F_Y \otimes F_Z(C).$$

To end the proof it is enough to show that for every fixed $\delta > 0$,

$$F_Y \otimes F_Z(C) \to 0 \quad \text{as } \varepsilon \to 0. \tag{2.3}$$

Notice that since $h(\cdot)$ is uniformly continuous, there is an $\eta = \eta(\delta)$ such that $|h(y + \varepsilon z - \varepsilon u) - h(y)| < \delta$ except when either $|z| > \eta/|\varepsilon|$ or $|u| > \eta/|\varepsilon|$. In particular, if $(y, z) \in C$ then either $|z| > \eta/|\varepsilon|$ or

$$(y, z) \in D = \left\{ (y, z) \colon \int_{|u| > \eta/|\varepsilon|} |h(y + \varepsilon z - \varepsilon u) - h(y)| F_Z(\mathrm{d}u) > \delta \right\}.$$

Therefore,

$$F_Y \otimes F_Z(C) \leqslant P(|Z| > \eta/|\varepsilon|) + F_Y \otimes F_Z(D).$$

Clearly, $P(|Z| > \eta/|\varepsilon|) \to 0$ as $\varepsilon \to 0$. The second term tends to zero, too. Indeed, $F_Y \otimes F_Z$ is absolutely continuous with respect to $\lambda \otimes F_Z$ and by Chebyshev's inequality,

$$\begin{aligned} \lambda \otimes F_Z(D) &\leqslant \frac{1}{\delta} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{|u| > \eta/|\varepsilon|} [h(y + \varepsilon z - \varepsilon u) + h(y)] F_Z(\mathrm{d}u) \lambda(\mathrm{d}y) F_Z(\mathrm{d}z) \\ &= \frac{2}{\delta} P(|Z| > \eta/|\varepsilon|) \to 0 \quad \text{as } \varepsilon \to 0. \end{aligned}$$

This establishes (2.3) and ends the proof of Theorem 1.1. □

**Lemma 2.1.** *Suppose $\mathscr{F}_\varepsilon \subset \mathscr{F}$ are σ-fields. Assume for every $\delta > 0$ there is an $\varepsilon_0(\delta) > 0$ and an $A \in \mathscr{F}$ with $P(A) > 1 - \delta$ such that for all $0 < |\varepsilon| < \varepsilon_0(\delta)$,*

$$\{S: S = B \cap A,\ B \in \mathscr{F}_\varepsilon\} \supset \{S: S = C \cap A,\ C \in \mathscr{F}\}. \tag{2.4}$$

*Then $E\{X \mid \mathscr{F}_\varepsilon\} \to X$ in $L_1$ as $\varepsilon \to 0$ for each integrable $X$.*

**Remark.** Obviously, (2.4) implies the equality of sets.

**Proof of Lemma 2.1.** By the standard approximation argument it is enough to consider $X = I_C$, the indicator function of an event $C \in \mathscr{F}$. Clearly,

$$E\{|E(I_C \mid \mathscr{F}_\varepsilon) - I_C|\} \leqslant E\{|E\{I_{C \cap A} \mid \mathscr{F}_\varepsilon\} - I_{C \cap A}|\} + 2P(A^c).$$

By (2.4) we have $E\{I_{C \cap A} \mid \mathscr{F}_\varepsilon\} - I_{C \cap A} = I_B\{E(I_A \mid \mathscr{F}_\varepsilon) - I_A\}$, for some $B \in \mathscr{F}_\varepsilon$. Therefore

$$E\{|E\{I_{C \cap A} \mid \mathscr{F}_\varepsilon\} - I_{C \cap A}|\} \leqslant 2E\{|I_A - P(A)|\} = 4P(A)P(A^c) < 4\delta.$$

This shows that $\| E\{I_C \mid \mathscr{F}_\varepsilon\} - I_C \|_1 \leqslant 6\delta$ for all $|\varepsilon| < \varepsilon_0(\delta)$. □

**Proof of Theorem 1.2.** Given $\delta > 0$ pick a measurable set $A \subset \Omega$ with $P(A) > 1 - \delta$ and such that the following two conditions are satisfied:

(i) $YI_A$ has finitely many values;

(ii) $ZI_A$ is bounded.

Put $\mathscr{F}_\varepsilon = \sigma(Y + \varepsilon Z)$. Clearly, for all $\varepsilon \neq 0$ such that $|\varepsilon|$ is small enough we have

$$\sigma(YI_A + \varepsilon ZI_A) = \sigma(YI_A, ZI_A). \tag{2.5}$$

This shows that condition (2.4) holds with $\mathscr{F}_\varepsilon$ given above and $\mathscr{F} = \sigma(Y, Z)$. Indeed, let $C \in \mathscr{F}$, i.e., $C = \{\omega: (Y, Z) \in \mathscr{V}\}$ for some Borel set $\mathscr{V} \subset \mathbb{R}^2$. Then

$$C \cap A = \{\omega: (YI_A, ZI_A) \in \mathscr{V}\} \cap A. \tag{2.6}$$

However, $\{\omega: (YI_A, ZI_A) \in \mathscr{V}\} = \{\omega: YI_A + \varepsilon ZI_A \in \mathscr{U}_\varepsilon\}$ for some Borel set $\mathscr{U}_\varepsilon \subset \mathbb{R}$, provided $\varepsilon \neq 0$ and $|\varepsilon|$ is small enough so that (2.5) holds. Therefore by (2.6),

$$C \cap A = A \cap \{\omega: YI_A + \varepsilon ZI_A \in \mathscr{U}_\varepsilon\} = A \cap \{\omega: Y + \varepsilon Z \in \mathscr{U}_\varepsilon\} = A \cap B$$

for some $B \in \mathscr{F}_\varepsilon$. By Lemma 2.1 the proof is complete. □

**Proof of Proposition 1.3.** Suppose $a \in \mathbb{R}$ is such that $P(Y = a) > 0$. Let $C = \{\omega: Y = a, Z = a\}$, $A_\varepsilon = \{\omega: Y + \varepsilon Z = (1 + \varepsilon)a\}$. Clearly, $P(C) > 0$ and $A_\varepsilon = C$ almost surely except for at most a countable number of $\varepsilon$'s. In particular, $P(A_\varepsilon) = P(C) > 0$.

Let $X = \phi(Z)$ for some $\phi$. By assumption $E\{\phi(Z) \mid Y + \varepsilon Z)\} \to E\{\phi(Z)\}$ in probability. On $A_\varepsilon$ we have

$$E\{\phi(Z) \mid Y + \varepsilon Z\} = \frac{1}{P(A_\varepsilon)} \int_{A_\varepsilon} \phi(Z)\, \mathrm{d}P = \frac{1}{P(C)} \int_C \phi(Z)\, \mathrm{d}P = \phi(a).$$

Therefore $E\{\phi(Z) \mid Y + \varepsilon Z\} = \phi(a)$ on a set of probability $P(C) > 0$. Since for each $\delta > 0$ we have

$$P\{|E\{\phi(Z) \mid Y + \varepsilon Z\} - E\{\phi(Z)\}| > \delta\} < P(C)$$

for all $\varepsilon \neq 0$ with $|\varepsilon|$ small enough, therefore $\phi(a) = E\{\phi(Z)\}$. The last equality, however, cannot be true for all functions $\phi$ unless $Z = a$ with probability one. This shows that if $Y$ is not continuous, then it is deterministic and ends the proof. □

## References

Szablowski, P.J. (1990), Expansions of $E\{X \mid Y + \varepsilon Z\}$ and their applications to the analysis of elliptically contoured measures, *Comput. Math. Appl.* **19**, 75–83.

Szablowski, P.J., (1986), On distributive relations involving conditional moments and the probability distributions of the conditioning random variable, *Demonstratio Math.* **19**, 721–746.