On the conditional expectation $E(X|X+W)$ in the case of independent random variables X, W

Ehrhard Behrends

Abstract. Let X, W be independent real valued random variables with finite expectation and $E(W) = 0$. We prove that only in the case $W = 0$ the conditional expectation $E(X|X+W)$ coincides with $X + W$.

The result is a consequence of the following cancellation theorem: Let P, Q, R be Borel probability measures on the real line such that the support of Q resp. R is contained in $\{x \leq 0\}$ resp. $\{x \geq 0\}$; then $P * Q = P * R$ implies that $Q = R (= \delta_0)$.

AMS-classification: 60E05;

keywords: conditional expectation, cancellation, convolution equation.

1. The problem

Let X and W be as in the abstract, we consider $X + W$ as a measurement of X in the presence of the "noise" W. If nothing is known about the specific situation one might naively assume that because of $E(W) = 0$ the best guess of X given $X + W$ could be $X + W$, i. e. that $E(X|X+W) = X + W$ (a. e.) should hold. However, a moment's reflection shows that under mild conditions on X this cannot happen unless $W = 0$.

> Suppose, e. g., that X, W are not only integrable but even L^2-functions. We assume that (without loss of generality) X has expectation zero. Then X and W are orthogonal in L^2, and we also know that $W \perp X + W$. Hence $W = 0$ by the Pythagorean law.

But what about the general situation? Is it possible that for independent X, W with $W \neq 0$ one can have $E(X|X+W) = X + W$? (In the language of [3]: Can it be that $X + W$ is self-consistent for X?) This equation makes sense if X is integrable, and if it holds W has to be integrable as well with $E(W) = 0$. And if both X, W are integrable the condition under consideration is equivalent with $E(W|X+W) = 0$, an equation which makes sense even if X is not necessarily integrable. Thus we arrive at the

Problem: Is $W = 0$ the only integrable random variable such that there is a (not necessarily integrable) X for which X, W are independent, and $E(W|X+W) = 0$?

Our main result will be that the answer is in the affirmative. It is proved in *section 2* where at first we investigate how the problem is related with *cancellation in convolution equations* involving positive measures. In general, for such measures P, Q, R, $P * Q = P * R$ does *not* imply that $Q = R$. In the case under consideration, however, this is admissible, since the support condition described in the abstract is satisfied. Our proof uses Fourier transformation, essential will be the fact that no (infinite) convex combination of the functions $\sin x, \sin 2x, \sin 3x, \cdots$ vanishes identically in a neighbourhood of zero.

2. Cancellation in convolution equations

Let X and W be independent real valued random variables such that W is integrable with expectation zero. Denote by P and S the distribution of X and W, respectively. We equip $\mathbb{R}^2$ with the product measure $P \otimes S$, and we may and will assume that the random variables under discussion are the coordinate projections. Now let $B \subset \mathbb{R}$ be a Borel set, and $\tilde{B} := \{(x,y) | x+y \in B\}$. The L^1-function $E(W|X+W)$ is measurable with respect to the σ-algebra consisting of all sets $\tilde{B}$, and $\int_{\tilde{B}} E(W|X+W) d(P \otimes S) = \int_{\tilde{B}} W d(P \otimes S) = \int_{\tilde{B}} y d(P \otimes S)$ for all $\tilde{B}$. Define a signed measure T on $\mathbb{R}$ by $T(B) := \int_B y dS$. Then, by the definition of convolution and the change-of-variable formula we obtain $\int_{\tilde{B}} E(W|X+W) d(P \otimes S) = (P * T)(B)$ (cf. [1], section 16). Consequently $E(W|X+W)$ vanishes iff $P * T$ is the zero measure. But $W = 0$ holds iff $T = 0$ so that the problem formulated in section 1 is the question whether or not P can be canceled in $P * T = 0$. If we decompose T canonically as $T = Q - R$ with positive measures we might also ask: Does $Q = R$ follow from $P * Q = P * R$? Note that the only information we have is that the support of Q (resp. R) lies in $\{x \geq 0\}$ (resp. $\{x \leq 0\}$).

Lets address first to a setting which is a little bit more general. For the study of (cancellation properties for) convolution equations the theory of topological groups is appropriate, and to define a convolution of two positive measures it suffices to assume that only one is finite. To illustrate what can happen consider the following examples:

- Consider on the group $\mathbb{Z}_4$ the measures $P := \delta_0 + \delta_1$, $Q := \delta_0 + \delta_2$, $R := \delta_1 + \delta_3$; then $Q \neq R$, but $P * Q = P * R$. (Here and in the sequel δ_a denotes the *Dirac measure* associated with a.)

- This time we consider measures on $\mathbb{Z}$: $P := \sum_{n \in \mathbb{Z}} \delta_{2n}$, $Q := \delta_{-2} + \delta_{-1}$, $T := \delta_1 + \delta_2$. Again $Q \neq R$ and $P * Q = P * R$.

- Let f be a nontrivial sufficiently smooth function from $\mathbb{R}$ to $\mathbb{R}$ which is 2π-periodic, vanishes on a suitable interval $[-\varepsilon, \varepsilon]$ and satisfies $f(-x) = -f(x)$ for all x. We can then write f as $f(x) = \sum_n a_n \sin(nx)$ for suitable real $a_1, a_2, \ldots$ with $\sum |a_n| < \infty$. Use these $a_1, a_2, \ldots$ to define positive finite measures P, Q on $\mathbb{Z}$ such that $\hat{R} - \hat{Q}$, the difference of the Fourier transforms, has the value $if(x)$ at any point $\exp(ix)$ (recall that, e. g., $\hat{R}(z) := \sum R(\{n\}) z^n$ for $|z| = 1$) .

  Now let g be a symmetric function which vanishes on $[-\varepsilon, \varepsilon]$, which is decreasing on $\{x \geq 0\}$ and which is also nontrivial and sufficiently smooth. Then g can be written as a series $b_0 + b_1 \cos x + b_2 \cos(2x) + \cdots$ with nonnegative $b_0, b_1, \cdots$ such that $0 < b_0 + b_1 + \cdots < \infty$. Define $P := 0.5 \sum_{n=0,1,\ldots} b_n(\delta_{-n} + \delta_n)$.

  Then $\hat{P}(\exp(ix) = g(x)$, and we conclude that $\hat{P}(\hat{R} - \hat{Q}) = 0$. Hence $P * Q = P * R$, but $Q \neq R$ (since $f \neq 0$).

The moral of the story is at follows:
1. Cancellation is not to be expected if no restriction on P, Q, R is imposed, even if the measures are finite (examples 1 and 3).
2. It seems to be that the possibility of "separating" the supports of P and Q and the finiteness of all measures is important: no cancellation if "separation" by some order relation cannot reasonably defined like in example 1, and also not in the case of nicely separated supports if unbounded measures are allowed (example 2).

In the case which is of importance here the situation is fortunately more favourable:

Proposition 2.1 *Let Q, R be finite positive Borel measures on the real line and P a signed measure on $\mathbb{R}$ with $P(\mathbb{R}) \neq 0$ such that $P * Q = P * R$. If the support of Q (resp. R) lies in $\{x \leq 0\}$ (resp. $\{x \geq 0\}$), then $Q = R$.*

Proof. Lets first restrict ourselves to the case when all measures are supported by the integers. The convolution condition implies that the Fourier transforms satisfy $\hat{P}\hat{Q} = \hat{P}\hat{R}$. But $\hat{P}(1) \neq 0$ so that $\hat{Q} = \hat{R}$ on a neighbourhood of 1. How can one conclude that $\hat{Q}(z) = \hat{R}(z)$ for *all* z with $|z| = 1$? The problem is that $\hat{Q}, \hat{R}$ are not necessarily analytic on a neighbourhood of $\{|z| = 1\}$. A key role in our argument will play the following

Claim: Let $\gamma_1, \gamma_2, \ldots$ be nonnegative numbers with $0 < \gamma_1 + \gamma_2 + \cdots < \infty$. Then there are arbitrarily small x such that $f_\gamma(x) := \sum \gamma_n \sin(nx) \neq 0$.

Proof of the claim: Let $\varepsilon > 0$ be given and n_0 be a number such that $\gamma_{n_0} > 0$. Choose $\delta > 0$ with $\delta \leq \varepsilon$ such that $\int_0^\delta \sin(n_0 x)dx > 0$. Since $\int_0^\delta \sin(nx)dx \geq 0$ for all n it follows that $\int_0^\delta f_\gamma(x)dx > 0$, hence there is an x in $[0, \delta]$ with $f_\gamma(x) > 0$.

Now suppose that $P \neq Q$ so that not both Q and R have their support in $\{0\}$. Then the numbers $\gamma_n := R(\{n\}) + Q(\{-n\})$ satisfy the assumption of the claim. Hence the imaginary part of $(\hat{R} - \hat{Q})(\exp(ix))$, which is just f_γ, vanishes on no neighbourhood of 0, and this shows that $P * Q \neq P * R$.

Only minor technical modifications are necessary to pass from $\mathbb{Z}$ to $\mathbb{R}$. The appropriate counterpart of the above claim reads as follows: Let G be a finite positive measure with support in $[0, \infty[$ such that $G(]0, \infty[> 0$; then $f_G(x) := \int_0^\infty \sin(tx)G(dt)$ vanishes on no neighbourhood of zero. The details are left to the reader. $\square$

By the results from the beginning of this section we conclude that the problem formulated in the introduction has a positive solution:

Proposition 2.2 *Let X, W be independent random variables such that W is integrable. Then $E(W|X + W) = 0$ implies that $W = 0$.*

We conclude with some *remarks*:

1. Any Borel measure M on $\mathbb{R}^r$ induces a Borel measure $\tilde{M}$ on $\mathbb{R}$ by $\tilde{M}(B) := M(B \times \mathbb{R}^{r-1})$, and $M \mapsto \tilde{M}$ behaves well with respect to convolution. This allows us to generalize proposition 2.1 by induction on r to Borel measures on $\mathbb{R}^r$; such a generalization could be phrased as follows: There are no Borel probability measures P, Q, R on $\mathbb{R}^r$ with $Q(\{(x_1, \ldots, x_r)|x_1, \ldots, x_r < 0\}) = 1 = R(\{(x_1, \ldots, x_r)|x_1, \ldots, x_r > 0\})$ such that $P * Q = P * R$.

2. Proposition 2.1 can be considered as a special case of assertions of the following type: Let P, Q, R be Borel probability measures on a topological

group G such that $P * Q = P * R$; if the support of P (resp. Q, R) lies in S_P(resp. S_Q, S_R), then one can conclude that $Q = R$.

It would be interesting to have more (non-obvious) examples of sets S_P, S_Q, S_R – in particular further examples with $S_P = G$ – where this holds. For example, is it in the case $G = \mathbb{Z}$ possible to take $S_P = \mathbb{Z}$ and any S_Q, S_R such that there is an integer $r \geq 0$ with $S_Q \subset]-\infty, r], S_R \subset [-r, \infty[$? (This is essentially the following analytical problem: Let $N \subset \mathbb{N}$ be finite, $\gamma_1, \gamma_2, \ldots \geq 0$ with $0 < \sum_{n \in \mathbb{N}} \gamma_n < \infty$; is it possible that the functions $\sum_{n \in N} \gamma_n \sin(nx)$, $\sum_{n \notin N} \gamma_n \sin(nx)$ coincide on a neighbourhood of zero?)

3. Let $\varepsilon > 0$ and C the space of real valued continuous functions on $[0, \varepsilon]$, provided with the supremum norm. By $x_n \in C$ we denote the function $t \mapsto \sin(nt)$. Consider $\hat{K} :=$ the collection of all $\sum \lambda_n x_n$, where $\lambda_1, \lambda_2, \ldots \geq 0, \sum \lambda_n = 1$. Then the essential step in the proof of proposition 2.1 was to show that $0 \notin \hat{K}$, and we proved this by providing for every n_0 a continuous linear functional y' such that $y'(x_n) \geq 0$ for all n, and $y'(x_{n_0})) > 0$. To phrase it otherwise: 0 lies not in $\hat{K}$ since there is an x' such that $x'(x_{n)} > 0$ for all n. This should be compared with the fact that 0 lies not in $K^- :=$ the closed convex hull of $x_1, \ldots$, if (and – by the Hahn-Banach separation theorem – even only if) x' can be chosen such that $\inf x'(x_n) > 0$.

We want to note that in contrast to the case of K^- the above does *not* characterize $\hat{K}$: There are Banach spaces and $x_1, x_2, \ldots$ such that $0 \notin \hat{K}$, but there is no x' which is strictly positive on all x_n (take e. g. in c_0 the vectors $x_1 := e_1, x_n := e_n - e_1$ for $n \geq 2$, where $e_1, \ldots$ denote the usual unit vectors). A simple description of $\hat{K}$ by functionals seems not to be possible.

(For a general discussion of $\hat{K}$, in particular its interesting topological properties, the reader is referred to section 22 in [2]).

REFERENCES

[1] P. BILLINGSLEY. *Probability and measure.* Wiley Serise in Probability, 3. edition, 1995.

[2] G. JAMESON. *Topology and normed spaces.* Chapman and Hall, 1974.

[3] TH. TARPEY, B. FLURY. *Self-consistency: A fundamental concept in statistics.* Statistical Science **11** (1996), 229.243.

I. Mathematisches Institut, Freie Universität Berlin, Arnimallee 2–6,
D-14195 Berlin, Germany; e-mail: behrends@math.fu-berlin.de