MAXIMUM VARIATION OF TOTAL RISK

Robin Pemantle [1,2]

Abstract: Let $Z > 0$ be a random time. The total risk of discovering Z in the next time interval $(t, t + dt)$ is never more variable than an exponential of mean one, which is achieved when the information up to time t is $\sigma(Z \wedge t)$.

1 Results

Scenario 1: You have a life insurance policy for one million dollars. The mortality tables for the entire population tell you that a lifespan of n years has probability q_n. Your premium for year n is $\$1,000,000 \cdot h_n$, where $h_n = q_n / \sum_{k=n}^{\infty} q_k$. In the absence of further information this is fair: you may choose each year whether to renew your policy, and your expected gain is always zero. If further information becomes available each year, the fair premium becomes $\$1,000,000 \cdot Q_n$, where Q_n is the conditional probability of dying in year n given all the information up to that point. How does the extra information affect the distribution of the lifetime total you pay for your policy?

Scenario 2: Random variables $\{X(e)\}$ are assigned to the edges of a graph. These values determine a random subset S of edges, called pivotal bonds. You know that $|S| \leq K$ with high

[1]Research supported in part by National Science Foundation grant # DMS 9300191, by a Sloan Foundation Fellowship, and by a Presidential Faculty Fellowship

[2]Department of Mathematics, University of Wisconsin-Madison, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706

probability. You order the edges $e(1), e(2), \ldots$ and look at the values $X(e_j)$ one at a time. You are interested (for reasons explained below) in the distribution of the random variable

$$W = \sum_{j=1}^{\infty} \mathbf{P}(e(j) \in A \mid X(e(i)) : i < j).$$

What bound can you get on $\mathbf{P}(W > \lambda K)$?

The purpose of this note is to prove an inequality that answers the questions in the two scenarios. The relevant notion of variability turns out to be the following one. Define a partial order $\preceq$ among random variables by

$$Y \preceq X \text{ if and only if for every convex } \phi, \mathbf{E}\phi(Y) \leq \mathbf{E}\phi(X),$$

where both expectations may be infinite. For $X \in L^1$, this is equivalent to the existence of $Y' \stackrel{\mathcal{D}}{=} Y$ and $X' \stackrel{\mathcal{D}}{=} X$ with $Y = \mathbf{E}(X \mid \mathcal{G})$ for some σ-field $\mathcal{G}$. Also, if $X \in L^1$ then $Y \preceq X$ is equivalent to the conjunction: $\mathbf{E}Y = \mathbf{E}X$ and $\mathbf{E}(Y - \lambda)^+ \leq \mathbf{E}(X - \lambda)^+$ for all real λ. [To see that $Y \preceq X$ implies $\mathbf{E}Y = \mathbf{E}X$, let ϕ be linear. Now assuming $\mathbf{E}Y = \mathbf{E}X$, it suffices in showing $Y \preceq X$ to consider convex ϕ with bounded derivative. Since $\mathbf{E}Y = \mathbf{E}X$, we may add a linear term and assume ϕ is increasing. Such a ϕ may be written as $\int (x - \lambda)^+ dF(\lambda)$.] The main result of this note is as follows. Let $\mathcal{E}$ denote an exponential random variable of mean one.

Theorem 1 (discrete case) *Let Z be a random positive integer and $\{\mathcal{F}_n\}$ be an increasing sequence of σ-fields. Let*

$$Y = \sum_{n=0}^{\infty} \mathbf{P}(Z = n + 1 \mid \mathcal{F}_n).$$

Then $Y \preceq \mathcal{E}$.

In order both to facilitate the proof and to accommodate future applications, I will pass to a rather general, continuous-time setting.

Theorem 2 (continuous case) *Let $A(t, \omega)$ be a random nondecreasing right-continuous function with $A(0) = 0$ and $A(\infty) = 1$, and let $\{\mathcal{F}_t\}$ be an increasing, right-continuous family of σ-fields. Let $\{A_t^p\}$ be the dual previsible projection of $\{A_t\}$ and let $R = A_\infty^p$. Then $R \preceq \mathcal{E}$.*

Before proving this, let me discuss the relation between the two theorems and the example scenarios. It is clear how Theorem 1 pertains to the insurance scenario. To see that the upper bound in variability is sharp, consider the continuous time insurance problem. Suppose that one's lifetime, Z, is a positive real random variable. The total risk is

$$R = \int_0^\infty \mathbf{P}(Z \in (t, t+dt) \,|\, \mathcal{F}_t) \,, \tag{1}$$

provided the RHS makes sense. Making sense of the RHS is where the dual previsible projection comes in. Let $A_t = \mathbf{1}_{[Z,\infty)}(t)$. The dual previsible projection of the increasing, right-continuous process $\{A_t\}$ formalizes the RHS of (1); see [4, Section VI.22] for further explanation. In the case where Z has a density f and $\mathcal{F}_t$ is the natural σ-field $\sigma(Z \wedge t)$, this turns into the familiar

$$R = \int_0^\infty \frac{f(t)}{1 - F(t)} \mathbf{1}_{Z>t} dt,$$

where $F(t) = \int_0^t f(s) ds$. It is well known that this has a mean-one exponential distribution independent of f. In fact this is true under much more general conditions, for instance when Z is a totally inaccessible stopping time and $\mathcal{F}_t$ is its natural filtration (see [2, prop. 3.28]). Two cases where the variability is less are the extreme cases: (1) $\mathcal{F}_t$ is trivial for all t, so $R = \int_0^\infty \mathbf{P}(Z \in (t, t+dt)) \equiv 1$; and (2) $\mathcal{F}_t = \sigma(Z)$ for all t, in which case $R = \int_0^\infty d\mathbf{1}_{t \leq Z} \equiv 1$ again. In general, the insurance company will be happy to know that the variance of the total premium of a policy based on up-to-date information will be less than the (easily computable) variance based on no updated information.

For the second scenario, let $\{s_1, \ldots, s_r\}$ be an ordering of the random set S, where $r \leq K$ is a random variable. For $1 \leq j \leq K$, let

$$Z_j = i \text{ if } e(i) = s_j$$

and $Z_j = \infty$ if $j > r$. Let

$$Y_j = \mathbf{P}(Z_j = \infty \,|\, \mathcal{F}_\infty) + \sum_{0 \leq n < \infty} \mathbf{P}(Z_j = n \,|\, \mathcal{F}_{n-1}).$$

It is easy to see that Theorem 1 extends to show that $Y_j \preceq \mathcal{E}$ for all j. Let Y_j' be the same as Y_j but without the term $\mathbf{P}(Z_j = \infty \,|\, \mathcal{F}_\infty)$, and let $W' = W\mathbf{1}_{|S| \leq K}$. Then

$$W' \leq \sum_{j=1}^{K} Y_j' \leq \sum_{j=1}^{K} Y_j \preceq K\mathcal{E}. \tag{2}$$

Thus, by an easy calculation, $\mathbf{P}(W' > \lambda K) \leq e^{1-\lambda}$. (The inequality (2) may also be derived directly from Theorem 2.)

The pivotal bond version of the problem comes from a paper of H. Kesten on first-passage percolation, [3]. Here, the method of bounded differences (an Azuma type inequality, c.f. Wehr and Aizenman [5]) is used to bound the variability of a first-passage time in terms of a conditional square function that turns out to be of the form discussed above. Kesten [3, Theorem 3] isolates the part of the argument that requires an upper tail bound on the conditional square function. Steps 2 and 3 of the Kesten's proof [3, Section 5] may be replaced by the result

$$\mathbf{P}\left[\sum_{k=1}^{N} \mathbf{E}(U_k \,|\, \mathcal{F}_{k-1}) \geq R, \sum_{k=1}^{N} U_k \leq T\right] \leq e^{1-R/T},$$

gotten by applying Theorem 2 to $T^{-1} \sum U_k \mathbf{1}_{\sum U_k \leq T}$.

Finally, to see that Theorem 1 is a special case of Theorem 2, begin with the hypotheses of Theorem 1 and let $A_t = \mathbf{1}_{Z \leq t+1/2}$. Let $\mathcal{F}_t' = \mathcal{F}_{\lfloor t \rfloor}$. Then $\{\mathcal{F}_t'\}$ is right-continuous and applying Theorem 2 gives

$$Y = \sum_{t+\frac{1}{2} \in \mathbf{Z}^+} \mathbf{P}(Z = t + \frac{1}{2} \,|\, \mathcal{F}_{t-\frac{1}{2}}) = A_\infty^p \preceq \mathcal{E}.$$

2 Proofs

Let $\{A_t^p\}$ be the dual previsible projection of $\{A_t\}$ as before, and let $\{{}^o\!A_t\}$ be the optional projection of $\{A_t\}$; the optional projection is a cádlág process such that for each t, ${}^o\!A_t$ is a version of $\mathbf{E}(A_t \,|\, \mathcal{F}_t)$.

Lemma 3 *The optional process $\{M_t\}$ defined by*

$$M_t = e^{A_t^p}(1 - {}^o\!A_t)$$

is a supermartingale with respect to $\{\mathcal{F}_t\}$.

The intuition behind this is pretty clear: $A_{t+dt}^p - A_t^p$ is the expected value of ${}^o\!A_{t+dt} - {}^o\!A_t$, so the total expected increase is $M_t(dA_t^p) - e^{A_t^p}\mathbf{E}(d{}^o\!A_t)$ which is never greater than zero. The proof is based on the following formula:

$$\begin{aligned}
M_t - M_s \;=\;& \int_s^t e^{A_{r-}^p}(1 - {}^o\!A_{r-})\, dA_r^p - \int_s^t e^{A_{r-}^p}\, d{}^o\!A_r \\
&+ \sum_{s<r\le t} \Big[e^{A_r^p}(1 - {}^o\!A_r) - e^{A_{r-}^p}(1 - {}^o\!A_{r-}) \\
&- e^{A_{r-}^p}(1 - {}^o\!A_{r-})(A_r^p - A_{r-}^p) + e^{A_{r-}^p}({}^o\!A_r - {}^o\!A_{r-}) \Big] \,. \qquad (3)
\end{aligned}$$

This formula may be derived from [1, page 334-335] by the following observation: since A is increasing, ${}^o\!A$ is a submartingale and A^p is increasing; then by [1, VIII (19.3)], the square bracket terms in [1, VIII (27.1)] vanish, resulting in (3). A more direct derivation without using the full strength of the stochastic Itô formula [1, VIII (27.1)] is possible. Observe for later use that $A_t^p - {}^o\!A_t$ is a martingale: assuming without loss of generality that $A_0 = 0$, one has for any stopping time T,

$$\mathbf{E}({}^o\!A_T - A_T^p) = \mathbf{E}(A_T - \int \mathbf{1}_{[0,T]}\, dA_s^p)\,;$$

since $\mathbf{1}_{[0,T]}$ is left continuous and adapted, it is predictable, and hence this becomes $\mathbf{E}(A_T - \int \mathbf{1}_{[0,T]}\, dA_s) = 0$.

Proof of Lemma 3: Rewrite the two integral terms in (3) as

$$\int_s^t (-{}^o\!A_{r-}) e^{A_{r-}^p}\, dA_r^p + \int_s^t e^{A_{r-}^p}\, d(A_r^p - {}^o\!A_r).$$

Using $\Delta {}^o\!A_r$ (respectively ΔA_r^p) to denote ${}^o\!A_r - {}^o\!A_{r-}$ (respectively $A_r^p - A_{r-}^p$), rewrite the summation as

$$\sum_{s<r\le t} e^{A_{r-}^p} \Big[e^{\Delta A_r^p}(1 - {}^o\!A_r) - (1 - {}^o\!A_{r-}) - (1 - {}^o\!A_{r-})\Delta A_r^p + \Delta {}^o\!A_r \Big] .$$

Combining the first, second and fourth terms inside the square brackets yields

$$(e^{\Delta A_r^p} - 1)(1 - {}^o\!A_r),$$

while expanding the third term yields

$$-(1 - {}^o\!A_r)\Delta A_r^p - (\Delta A_r^p)^2 + \Delta A_r^p(\Delta A_r^p - \Delta {}^o\!A_r).$$

The quantity in square brackets may therefore be rewritten as

$$(e^{\Delta A_r^p} - 1 - \Delta A_r^p)(1 - {}^o\!A_r) - (\Delta A_r^p)^2 + \Delta A_r^p(\Delta A_r^p - \Delta {}^o\!A_r)$$

and equation (3) now becomes

$$\begin{aligned} M_t - M_s \;=\; & \int_s^t (-{}^o\!A_{r-}) e^{A_{r-}^p}\, dA_r^p + \int_s^t e^{A_{r-}^p}\, d(A_r^p - {}^o\!A_r) \\ & + \sum_{s<r\le t} e^{A_{r-}^p}(-{}^o\!A_r)(e^{\Delta A_r^p} - 1 - \Delta A_r^p) + \sum_{s<r\le t} e^{A_{r-}^p}(e^{\Delta A_r^p} - 1 - \Delta A_r^p - (\Delta A_r^p)^2) \\ & + \sum_{s<r\le t} e^{A_{r-}^p} \Delta A_r^p (\Delta A_r^p - \Delta {}^o\!A_r). \end{aligned}$$

The conditional expectations given $\mathcal{F}_s$ may be seen to be nonpositive term by term. The first integral is everywhere nonpositive. The second is the integral of a previsible process against a martingale and hence has zero expectation given $\mathcal{F}_s$. The first summation is everywhere nonpositive, as is the second, since $\Delta A_r^p = \mathbf{E}(\Delta A_r \mid \mathcal{F}_{r-}) \in [0,1]$ for all r, and $e^z \le 1 + z + z^2$ for $z \in [0,1]$. Finally, the third summation is the integral of the previsible process $e^{A_{r-}^p}\Delta A_r^p$ against the martingale $A_r^p - {}^o\!A_r$, and therefore has zero conditional expectation given $\mathcal{F}_s$. Thus M_t is a supermartingale.

Proof of Theorem 2: Fix a real $\lambda > 0$ and define a stopping time $\tau = \inf\{t \ge 0 : A_t^p \ge \lambda\}$. The purpose of the argument between here and (4) is to handle the case where, due to jumps, $A_{\tau-}^p < \lambda$. If you are not worried about jumps, skip ahead to (4) and read only the first expression inside each subsequent expectation.

Since τ is previsible, there are times $\tau_n \ne \tau$ increasing to τ almost surely, and it follows that

$$\mathbf{E} M_{\tau-} = \mathbf{E} \lim M_{\tau_n} \le \liminf \mathbf{E} M_{\tau_n} \le \mathbf{E} M_0.$$

Now define a random variable X by setting $X = 0$ when $\tau = \infty$, setting $X = e^{\lambda}(1 - {}^o\!A_\tau)$ when $\Delta A^p_\tau = 0$, and otherwise setting

$$X = e^{\lambda}\left[\frac{A^p_\tau - \lambda}{A^p_\tau - A^p_{\tau-}}(1 - {}^o\!A_{\tau-}) + \frac{\lambda - A^p_{\tau-}}{A^p_\tau - A^p_{\tau-}}(1 - {}^o\!A_\tau)\right].$$

The following computation shows that $\mathbf{E}X < \mathbf{E}M_{\tau-} \leq 1$ in the case where $\Delta A^p_\tau \neq 0$. Make use of the facts that $A^p_\tau, A^p_{\tau-}$ and ${}^o\!A_{\tau-}$ are all in $\mathcal{F}_{\tau-}$, and $\mathbf{E}(\Delta {}^o\!A_\tau - \Delta A^p_\tau \,|\, \mathcal{F}_{\tau-}) = 0$ to write:

$$\begin{aligned} & \mathbf{E}(X - M_{\tau-} \,|\, \mathcal{F}_{\tau-}) \\ = \;& (e^{\lambda} - e^{A^p_{\tau-}})(1 - {}^o\!A_{\tau-}) - e^{\lambda}(\lambda - A^p_{\tau-}) \\ \leq \;& e^{\lambda} - e^{A^p_{\tau-}} - e^{\lambda}(\lambda - A^p_{\tau-}). \end{aligned}$$

This is less than or equal to 0 since $e^z - e^y - e^z(z - y) \leq 0$ for $z \geq y \geq 0$. In the cases $\Delta A^p_\tau = 0$ or $\tau = \infty$, the conclusion that $\mathbf{E}(X - M_{\tau-} \,|\, \mathcal{F}_{\tau-}) \leq 0$ still holds, some terms having dropped out of the above computation.

Combining this result with the fact that $\{M_t\}$ is a supermartingale shows that $e^{-\lambda} = e^{-\lambda}\mathbf{E}M_0 \geq e^{-\lambda}\mathbf{E}M_\tau \geq e^{-\lambda}\mathbf{E}X$. Thus

$$e^{-\lambda} \geq \mathbf{E}\left[(1 - {}^o\!A_\tau) + ({}^o\!A_\tau - {}^o\!A_{\tau-})\frac{A^p_\tau - \lambda}{A^p_\tau - A^p_{\tau-}}\right]. \tag{4}$$

Taking conditional expectations with respect to $\mathcal{F}_{\tau-}$ shows that ${}^o\!A$ may be replaced by A, yielding

$$e^{-\lambda} \geq \mathbf{E}\left[(1 - A_\tau) + (A_\tau - A_{\tau-})\frac{A^p_\tau - \lambda}{A^p_\tau - A^p_{\tau-}}\right].$$

Since $A_\infty = 1$, we may write the RHS as the stochastic integral

$$\mathbf{E}\int\left(\mathbf{1}_{(\tau,\infty)}(t) + \frac{A^p_\tau - \lambda}{A^p_\tau - A^p_{\tau-}}\mathbf{1}_{t=\tau}\right) dA_t.$$

The integrand is previsible, so this becomes

$$\begin{aligned} e^{-\lambda} \;&\geq\; \mathbf{E}\int\left(\mathbf{1}_{(\tau,\infty)}(t) + \frac{A^p_\tau - \lambda}{A^p_\tau - A^p_{\tau-}}\mathbf{1}_{t=\tau}\right) dA^p_t \\ &=\; \mathbf{E}(A^p_\infty - \lambda)^+. \end{aligned}$$

Thus the total risk R satisfies $\mathbf{E}(R-\lambda)^+ \leq e^{-\lambda}$ for every positive λ, which, along with the fact that $\mathbf{E}R = 1$, suffices to prove $R \preceq \mathcal{E}$. $\square$

References

[1] Dellacherie, C. and Meyer, P. (1982). Probabilities and Potential B: theory of martingales. Translated and prepared by J. P. Wilson. North-Holland mathematical studies volume 72, North-Holland: Amsterdam.

[2] Jeulin, T. (1980). Semi-martingales et grossissement d'une filtration. Lecture Notes in Mathematics # 833. Springer-Verlag: New York.

[3] Kesten, H. (1993). On the speed of convergence in first-passage percolation. *Ann. Appl. Prob.* **3** 296 - 338.

[4] Rogers, L.C.G. and Williams, D. (1987). Diffusions, Markov processes, and martingales. Volume II: Itô calculus. John Wiley & Sons: New York.

[5] Wehr, J. and Aizenman, M. (1990). Fluctuations of extensive functions of quenched random couplings. *J. Stat. Phys.* **60** 287 - 306.