Optimal Design, Robustness and Risk Aversion

M.E.J. Newman, Michelle Girvan and J. Doyne Farmer

Summarised: 7^{th} of March 2018
by Ida Torkjellsdatter Storehaug and Utkarsh Detha

In "Optimal Design, Robustness, and Risk Aversion"[3], Newman, Girvan and Doyne Farmer analytically derive the origin of power law behaviour in "highly optimized tolerance" (HOT) systems and suggest an alternative approach that incorporates risk aversion while preserving the desired qualities of the HOT model, fairly closely. Recent results show that the difference between multiple measurements of the same quantity, is consistent with a heavy-tailed Student-t distribution, meaning that a measurement more often than expected is in disagreement with the estimated value.[1] Incorporating risk aversion truncates the heavy tails, and could potentially lead to more consistent results when comparing experiments.

In their article, Newman et al. use a model of forest fires to show why failure of complex systems follow a power-law distribution. They argue that this is a trait of HOT systems – that a system designed for maximizing the outcome naturally organizes into statistically unlikely states that are robust to perturbations they were designed to handle, but fragile to rare perturbations. They hypothesize that rather than optimizing the yield of a system, one should optimize the utility, resulting in a model they decided to call COLD ("constrained optimization with limited deviations").

Like the proposers of the HOT model[2], the authors consider a HOT forest fire model. Unlike the proposers, they use a continuum model instead of a lattice based one. In this model a forester is charged with optimizing the system: finding the distribution of trees that will maximize the harvest. The area of the patches of trees is $s(\mathbf{r})$. However, occasional fires, ignited by sparks that arrive by a given distribution, $p(\mathbf{r})$, burn down patches of forest. Areas of trees are separated by firebreaks, making the system more robust. Still, any changes in the spark distribution or unforeseen events, can really lower the yield of the forest, an example of being robust to expected perturbations, but fragile to unforeseen ones.

The yield is $Y = 1 - s(\mathbf{r}) - F$, where F is the cost in terms of yield of constructing the firebreak around the patch m. $F = agds_m^{(d-1)/d}$, where a is the construction cost per unit length/surface, g is a geometric factor of order 1 that depends on the shape of the patch, d is the dimension and s_m is the value of $s(\mathbf{r})$ in patch m. The authors find the total area of the firebreaks by summing over all m, inserting this into the expression for Y:

$$Y = 1 - \int p(\mathbf{r}) s(\mathbf{r}) d^d r - agd \int s(\mathbf{r})^{-1/d} d^d r \tag{1}$$

By calculating the maximum yield ($\delta Y / \delta s(\mathbf{r}) = 0$) one derives an expression for $s(\mathbf{r})$:

$$s(\mathbf{r}) = \left[\frac{ag}{p(\mathbf{r})} \right]^{d/(d+1)} \tag{2}$$

The distribution $\rho(s)$ of fire sizes, when taking the optimal choice of patch sizes, is given as:

$$\begin{aligned}\rho(s) =& p(\mathbf{r})\frac{d^d r}{ds} = -ag\frac{d+1}{d}p(\mathbf{r})\frac{d^d r}{dp}s^{-(2+1/d)} \\ =& C \cdot p(\mathbf{r})\frac{d^d r}{dp} \cdot \text{"}powerlaw\text{"}\end{aligned} \tag{3}$$

where $p(\mathbf{r})$ is given by eq. (2). The power law behaviour becomes apparent from the term $s^{-(2+1/d)}$. However, to understand the behaviour of the terms containing $p(\mathbf{r})$, we need to assume a certain distribution – a normal distribution would be the obvious choice. Consider a spark distribution, $p(\mathbf{r})$, in two dimentions ($d = 2$) having the form of two normalized Gaussians with different widths, σ_x, σ_y. Plugging this distribution into eq. (3), one find the distribution of the event sizes to be $\rho(s) = 3\pi\sigma_x\sigma_y ags^{-5/2}$. Thus the model generates a power law with the slope of $-\frac{5}{2}$.

In their paper the authors argue that the optimization performed above is problematic, and can lead to ruinous outcomes. Instead one could strike a reasonable compromise between the optimization of the yield and the robustness, by following the proposed COLD model. The COLD forest fire model incorporate risk aversion, described by the utility function $u(s)$, which is a function of the loss s:

$$u(s) = \frac{(1-s)^\alpha}{\alpha} \tag{4}$$

Then by maximizing the average utility function $U = \int p(\mathbf{r})u(s(\mathbf{r}))d^d r$, for fixed F, we obtain the following expression:

$$\frac{dp}{ds} = \lambda\frac{(\alpha + 1/d)s - (1 + 1/d)}{(1-s)^\alpha s^{2+1/d}} \tag{5}$$

Substituting eq. (5) into eq. (3), instead of eq. (2), we find an expression for $\rho(s)$ for the COLD model, maximizing utility instead of yield. Simply by setting $\alpha = 1$, the HOT model is retrieved; For $\alpha < 1$ we have a risk-averse utility function (COLD). The authors then plot this function for variable values of α compared to the HOT model, and find that incorporating the utility function truncates the power law behaviour of the event size distribution. However, the distribution of utilities still follows a power law.

The authors find that you pay a cost for risk aversion – the harvest of the forester is smaller than with a HOT system. However, compared to HOT they find that the cost in terms of average system yield is only a few percent smaller, but the reduction of large losses is substantial. They also conclude that the suppression of the heavy tails will make the system more robust to other, untested fluctuations.

Their result has implications for scientific measurements in general. Scientific research, with people, hardware and software, are good examples of complex systems designed to optimize outcomes: the best measurements possible. Such experiments will be sensitive to unexpected events, resulting in a power law distribution of systematic errors, and the final measurement of such experiments are more likely to be outliers than expected. If applied to scientific research the results of Newman, Girvan and Doyne Farmer should tell us that researchers willing to accept a small loss in experiment performance could obtain more consistent results.

References

[1] David C. Bailey. Not Normal: the uncertainties of scientific measurements. 2016.

[2] J. M. Carlson and John Doyle. Highly optimized tolerance: Robustness and design in complex systems. *Phys. Rev. Lett.*, 84:2529–2532, Mar 2000.

[3] M. E. J. Newman, Michelle Girvan, and J. Doyne Farmer. Optimal design, robustness, and risk aversion. *Phys. Rev. Lett.*, 89:028301, Jun 2002.