Institute for Homeland Security Solutions
Applied research • Focused results

June 2009

Adversarial Risk Analysis: Decision Making When There Is Uncertainty During Conflict

Project Lead

David Banks, PhD, Duke University

Statement of Problem

Counterterrorism requires decision makers to allocate defensive resources in situations for which the kind of attack and the likely consequence of an attack are unknown. The two current tools for analyzing such problems are inadequate. Classical game theory (Gibbons, 1992; Myerson, 1991) assumes that the costs and benefits are known, does not use partial information (e.g., military intelligence), and produces results that humans find unrealistic (Bier & Cox, 2007). Statistical risk analysis assumes that the adversary is “nature” rather than an intelligent opponent who seeks to exploit weaknesses (Bedford & Cooke, 2001; Singpurwalla, 2006). Neither approach takes explicit account of resource constraints under which managers must operate (Brown, Carlyle, & Wood, 2008).

This research brief focuses on new strategies for repairing these deficiencies. In particular, it considers Bayesian versions of game theory (Harsanyi, 1967, 1968a, 1968b, 1982; Kadane & Larkey, 1982; Rios Insua, Rios, & Banks, in press). Some of these methods produce results that seem better suited to real-world problems than those that arise as minimax solutions (i.e., the standard game theory criterion that minimizes the worst that can happen, while providing no protection against the second-worst attack and ignoring prior history and relevant information). For example, in the Prisoner's Dilemma, both opponents are compelled by game theory to choose a bad solution that protects against the worst outcome, but in empirical studies, about half the time people cooperate to achieve a better solution.

The approach described in this brief applies beyond counterterrorism—corporate competition and federal regulation are two examples. The underlying theory is still being developed. This brief describes three illustrations of the approach using simple poker, an auction, and the 2003 decision on counterbioterrorism policy regarding smallpox.

Background

Camerer (2003) described, from a behavioral perspective, a number of deficiencies with classical game theory. These deficiencies include bounded rationality of the opponents, their computational limitations, and their penchant for unwarranted optimism. Bounded rationality refers to the fact that people are imperfect in their logic; they look ahead only a few steps in the game, they replace calculation with intuition, and they seek analyses that support decisions they prefer. Computational limitations imply that people do not perform the in-depth linear programming analysis needed to find the minimax solution. Such solutions are complex; each player must find the action that maximizes benefit under the constraint that the opponent is doing a symmetric analysis. Furthermore, minimax solutions are often disappointing—they represent extreme points in the decision space, so they tend to be both pessimistic and expensive, and. thus, policy makers are uncomfortable.

Sandler and Arce (2003) reviewed classical game theory in the context of counterterrorism. Their examples are necessarily simplistic, because it is difficult, and perhaps impossible, to scale-up minimax solutions to applications with thousands of possible actions, each with different costs and uncertain benefits. More strongly, game theory founders when the outcomes of the decisions are stochastic; this has led some agencies to attempt to substitute risk analysis for game theory (Parnell et al., 2008), but that replacement overlooks the adversarial nature of the problem. These concerns suggest that a combination of the two methods, as described in the following, may lead to better decision making.

To illustrate some of the issues, begin by considering two typically simple game theory problems: poker and auctions. For both cases, a Bayesian perspective and risk analysis of uncertain outcomes leads to reasonable solutions. Then consider a more sophisticated example pertaining to the analysis of the 2003 decision on smallpox vaccination policy.

First, consider the simplest possible game of poker: there are two opponents (Lisa and Bart), a fixed ante (a), and one round of betting with a fixed bid (b). Bellman and Blackwell (1949) worked out the game theory for this situation—the minimax solution never involves bluffing. If Lisa calculates that the probability that she is holding the best hand is greater than p^*, then she bids; otherwise, she folds. Bart has a similar rule for folding or checking; his rule is entirely determined by the same p^* (which depends upon a and b, the sizes of the ante and the fixed bid). Both players, being rational minimaxers, completely deduce their opponent's rule.

In contrast, a Bayesian analysis of this game finds an interestingly different solution (Banks, 2009). Lisa has a bluffing function $g(p)$ that determines her probability of bidding as a function of the strength (p) of her hand. She obtains that function by "mirroring" the analysis that Bart does of her decision problem, under the assumption that Bart believes that her bluffing function is $gB(p)$. Using Bayes theorem, Lisa then optimizes her bluffing function in response. If the $gB(p)$ is in fact the minimax rule (a step function at p^*), then Lisa's best play coincides with the classical result. But if Bart is not a rational minimaxer and Lisa has approximately mirrored his thinking, then she can exploit his vulnerability. In that case, her optimal bluffing function is a step function at p^{**}, where $p^{**} < p^*$.

In real life, players often have strong and accurate beliefs about their opponent's bluffing function, $g(p)$. The Bayesian approach enables human judgment to be used and yields solutions that resemble expert play. This parallels counterterrorism situations, where there is often useful but imperfect intelligence information, prior history, and so forth, that are available to decision makers.

The second illustration of this approach is an auction. Suppose that Bart and Lisa each submit a sealed bid on a comic book. Lisa does not know the personal value that Bart puts on the book (and conversely), but she knows her own value for the book (v^*) and her expected gain from a bid of v is $(v^* - v)$ P[v wins the bid]—this is her increase in value from a successful bid of v, times the probability that the bid is successful (Raiffa, 2002). She wants to maximize this, and needs to mirror Bart's analysis is order to estimate P[v wins the bid].

To explain Lisa's reasoning, some notation is needed:

- W^* is Bart's value for the book; it is unknown to Lisa and, thus, she places a subjective Bayes distribution H_L on it.
- v^* is the random variable that Lisa thinks Bart uses to describe her value for the book; she does not know this, and so places the distribution H_B on it.
- F is Lisa's belief about the distribution of Bart's bid.
- G is Lisa's inference about Bart's distribution on Lisa's bid.

This seems a bit convoluted but, in fact, it is a formal expression of common human thinking. Lisa has a belief (H_L) about how much Bart values the book. She also has a belief (H_B) about what Bart thinks the book is worth to her. Using that, she must solve to find (F), the distribution of Bart's bid, enabling her to calculate P[v wins the bid] and, thus, select her bid (v) to maximize her expected gain.

In classical game theory, this reasoning would go further, into an infinite regress, but this level of modeling is sufficient to handle second-order secret information. For example, one can imagine an auction in which there is secret information that the comic book was signed by Matt

Groening. Lisa's H_L reflects her belief about whether Bart knows of the signature; her H_B indicates whether she thinks that Bart knows that she knows of the signature. She could try to model whether Bart thinks that she thinks that he knows of the signature, and so forth, but that goes beyond standard human cognition (i.e., it is an example of bounded rationality [Gigerenzer & Selten, 2002]).

The solution to this auction problem is found by solving a system of distributional equations:

$$\text{argmax}_{\{w > 0\}} (W^* - w) G(w) \sim F$$

$$\text{argmax}_{\{v > 0\}} (V^* - v) F(v) \sim G.$$

This system has a fixed point, so standard methods enable Lisa to derive F (Banks, 2009).

The third illustration concerns the 2003 decision by the U.S. government on the kind of defense it would mount against a smallpox attack. There were four options: stockpile vaccine; stockpile and perform biosurveillance; stockpile, perform biosurveillance, and inoculate first responders; and inoculate everyone not medically contraindicated. These options responded to three possible scenarios: no smallpox attack, a minor attack (similar to the anthrax letters), and a large-scale attack.

Classical game theory would set this up as a 3×4 matrix whose entries are pairs of numbers that represent the costs to each opponent (the costs may be positive or negative; in a zero-sum game, the benefit to one opponent is the negative of the cost to the other). As shown in Table 1, the entries in the first row, first column are the cost to the United States (X_{11}) and the cost to the terrorist (Y_{11}) of stockpiling vaccine when no attack is made, and so forth for the other cells in the matrix. If one knew those values, then linear programming could find the Nash equilibrium, which would represent the optimal decision under classical game theory.

Table 1. Normal Form Cost Matrix for the 2002 Smallpox Decision

Scenario	Stockpile	Surveillance	Responder	Mass Program
No attack	(X_{11}, Y_{11})	(X_{12}, Y_{12})	(X_{13}, Y_{13})	(X_{14}, Y_{14})
Minor attack	(X_{21}, Y_{21})	(X_{22}, Y_{22})	(X_{23}, Y_{23})	(X_{24}, Y_{24})
Major attack	(X_{31}, Y_{31})	(X_{32}, Y_{32})	(X_{33}, Y_{33})	(X_{34}, Y_{34})

However, in practice, the actual costs in each cell are unknown. To assess these, one should use statistical risk analysis. In that case, expert opinion and historical data would be used to place a bivariate distribution over the joint costs to the defender and the attacker, conditional on the corresponding row and column of the matrix. (Expert opinion is often unreliable [O'Hagan et al., 2006], but it becomes better as uncertainties are reduced, and this use of the row and column information should sharpen the quality of the expert elicitation.)

As a result of these elicitations, one gets a multivariate distribution over the entire cost matrix. One does not know which matrix nature will use in a particular situation, but the distribution indicates which is more likely. At this point, some analysts replace the random costs in the matrix by their expected values and find the minimax solution (Brown, Carlyle, Salmeron, & Wood, 2006). But this process is incorrect; the best solution for average game is not the same as the solution that is, on average, the best (i.e., taking expectations and computing minimax equilibria do not commute).

Instead, Banks and Anderson (2006) considered two alternatives. The first generates random matrices according to the multivariate distribution, computes the minimax solution for each, and tallies the number of times that each of the defense strategies is the best. For a realistic multivariate distribution developed from extensive interviews with experts at the U.S. Food and Drug Administration, U.S. Department of Health and Human Services, U.S. Department of Homeland Security, BES, and other federal agencies, it turned out that each of the defenses under consideration was often a minimax solution and, thus, there was little useful guidance for decision makers.

Banks and Anderson's (2006) second analysis had experts place a Bayesian distribution over the actions of the enemy, conditional on information that would be publicly available and, hence, known to terrorists once one of the four options for defense against a smallpox attack had been chosen by the U.S. government. From that, one can calculate the decision that has minimum expected loss. In that analysis, the result was clear—it was best to stockpile vaccine but not pursue any further defensive investments (allowing money to be retargeted to other threats). In retrospect, based on the following 6 years of experience, it seems that this decision would have been the best choice among the four alternatives under consideration.

Synthesis

These examples show that Bayesian approaches to game theory lead to usefully different solutions than those obtained by classical game theory. The minimax solutions do not take reasonable account of uncertainty in the risk analyses, the value of intelligence reports and prior experience, and the possibility of modeling the decision processes of one's adversaries. But, Bayesian methods can handle these features of the problem very naturally.

In the context of counterterrorism, the smallpox example illustrates most clearly the potential for combining Bayesian game theory with statistical risk analysis. It appears that even an approximate Bayesian analysis produces guidance that is both reasonable and actionable.

Future Directions

This is a new area, and it is rich in research targets. Future work is needed in the following fields:

- the effect of resource constraints, so that decision makers can manage the trade-offs in different kinds of threat reduction, especially when there are partially overlapping defense actions (e.g., stockpiling smallpox vaccine protects well against the smallpox threat, but investing that money in counterintelligence may protect against many additional threats, though not as well) (Bier, 2007; Bier, Oliveros, & Samuelson, 2007)
- exploration of behavioral decision making by adversaries in the context of several extensions of the previous examples: poker games with shared knowledge, multiple players, and repeated rounds of betting; auctions with more than two bidders and collusion; auctions with alliance formation, in which each partner has a different utility function on the resources it can contribute and the benefits it receives
- analysis of other counterterrorism problems, such as optimal routing of convoys across networks that contain improvsied explosive devices (IEDs) using historical information and intelligence; optimal allocation of counterbioterrorism resources to multiple pathogens (e.g., smallpox, anthrax, influenza); hierarchical decision making, in which broad decisions are made by top executives, but successive refinements are made by managers with less authority

Contact Information

David Banks, Ph.D.
Department of Statistical Sciences
Box 90251
Duke University
Durham, NC 27708
Phone: 919-684-3743
Email: banks@stat.duke.edu

David Banks, Ph.D., is a professor in the Department of Statistical Sciences at Duke University. He is a past-chair of the American Statistical Association's Sections on (1) Risk Analysis and (2) Statistics in Defense and National Security. He is the editor-in-chief of the *Journal of the American Statistical Association*. He served on the National Academy panels for (1) Biodefense Analysis and Countermeasures and (2) Biological Agent Risk Analysis. His research expertise includes data mining and risk analysis. He is a member of the Board of

Directors of the American Statistical Association. David Banks has a Ph.D. in statistics from Virginia Tech and has held previous positions at the University of California, Berkeley; University of Cambridge, Carnegie Mellon University, the National Institute of Standards and Technology, the U.S. Department of Transportation, and the U.S. Food and Drug Administration.

References

Banks, D. (2009). Unpublished research.

Banks, D., & Anderson, S. (2006). Game theory and risk analysis in the context of the smallpox threat. In A. Wilson, G. Wilson, & D. Olwell (Eds.), *Statistical methods in counterterrorism* (pp. 9–22). New York: Springer.

Bedford, T., & Cooke, R. (2001). *Probabilistic risk analysis: Foundations and methods.* Cambridge, United Kingdom: Cambridge University Press.

Bellman, R., & Blackwell, D. (1949). Some two-person games involving bluffing. *Proceedings of the National Academy of Sciences USA*, *35*, 600–605.

Bier, V. (2007). Choosing what to protect. *Risk Analysis*, *27*, 607–620.

Bier, V., & Cox, Jr., L. A. (2007). Probabilistic risk analysis for engineered systems. In W. Edwards, D. von Winterfeldt, & R. Miles (Eds.), *Advances in decision analysis* (pp. 279–301). Cambridge, United Kingdom: Cambridge University Press.

Bier, V., Oliveros, S., & Samuelson, L. (2007). Choosing what to protect. *Journal of Economic Literature, 9*, 563–587.

Brown, G., Carlyle, M., Salmeron, J. & Wood, K. (2006). Defending critical infrastructure. *Interfaces, 36*, 530–544.

Brown, G., Carlyle, W. M., & Wood, R. (2008). Optimizing Department of Homeland Security Defense investments: Applying defender-attacker(-defender) optimization to terror risk assessment and mitigation. In *Report on Methodological Improvements to the Department of Homeland Security's Biological Agent Risk Analysis, Appendix E*. Washington, DC: National Academies Press.

Camerer, C. (2003). *Behavioral game theory.* Princeton, NJ: Princeton University Press.

Gibbons, R. (1992). A *primer in game theory.* Essex, England: Pearson Education Ltd., Harlow.

Gigerenzer, G., & Selten, R. (2002). *Bounded rationality*. Cambridge, MA: MIT Press.

Harsanyi, J. (1967). Games with incomplete information played by “Bayesian” players, I-III. Part I. The basic model. *Management Science*, *14*, 159–182.

Harsanyi, J. (1968a). Games with incomplete information played by “Bayesian” players, I-III. Part II. Bayesian equilibrium points. *Management Science, 14*, 320–334.

Harsanyi, J. (1968b). Games with incomplete information played by “Bayesian” players, I-III. Part III. The basic probability distribution of the game. *Management Science, 14*, 486–502.

Harsanyi, J. (1982). Subjective probability and the theory of games: Comments on Kadane and Larkey's paper. *Management Science, 28*(2), 120–124.

Kadane, J. B., & Larkey, P.D. (1982). Subjective probability and the theory of games. *Management Science*, *28*, 113–120; reply: 124.

Myerson, R. (1991). *Game theory: Analysis of conflict*. Cambridge, MA: Harvard University Press.

O'Hagan, A., Buck, C., Daneshkhah, A., Eiser, J., Garthwaite, P., Jenkinson, D., et al. (2006). *Uncertain judgments: Eliciting experts' probabilities*. New York: Wiley.

Parnell, G., Banks, D., Borio, L., Brown, G., Cox, L., Gannon, J., et al. (2008). *Report on methodological improvements to the Department of Homeland Security's biological agent risk analysis.* Washington, DC: National Academies Press.

Raiffa, H. (2002). *Negotiation analysis*. Cambridge, MA: Harvard University Press.

Rios Insua, D., Rios, J., & Banks, D. (in press). Adversarial risk analysis. *Journal of the American Statistical Association*.

Sandler, T., & Arce M., D. G. (2003). Terrorism and game theory. *Simulation and Gaming*, *34*, 319–337.

Singpurwalla, N. (2006). *Reliability and risk: A Bayesian perspectiv*e. Chichester, United Kingdom: Wiley.