David Pollard
Yale University
http://www.stat.yale.edu/~pollard/

WHAT IS RANDOMIZATION?
Talk at University of Chicago
7 February 2005

- Sample spaces are arbitrary. We change them at our convenience.
  We like to think of random variables as functions on Ω.
  We like expectations to follow rules for integrals with respect to countably additive measures.

- Model: $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ on $(\Omega, \mathcal{F})$.
  Statistic T taking values in $(\mathcal{T}, \mathcal{B})$. That is, T is $\mathcal{F}\backslash\mathcal{B}$-measurable.
  $\mathbb{Q}_\theta$ = distribution of T under $\mathbb{P}_\theta$.
- Say that T is sufficient for $\mathcal{P}$ if there is a randomization[t] not depending on θ, such that

$$t \sim \mathbb{Q}_\theta \qquad \text{and} \qquad \widetilde{\omega} \mid t \sim \text{randomization}[t] \qquad \text{imply} \qquad \widetilde{\omega} \sim \mathbb{P}_\theta$$

- Is randomization[t] given by a Markov kernel: a probability measure π_t on $\mathcal{F}$ for which $\pi_t\{\omega : T(\omega) = t\}$, almost surely [$\mathbb{Q}_\theta$]?
- Is randomization[t] given by a Kolmogorov conditional expectation?

  $\kappa(t, X)$ is a version of $\mathbb{E}_\theta(X \mid T = t)$ for which, with $\mathbb{Q}_\theta$ probability one for each θ:

$$\begin{aligned} &\kappa(t, 1) = 1 \\ &\kappa(t, \alpha_1 X_1 + \alpha_2 X_2) = \alpha_1 \kappa(t, X_1) + \alpha_2 \kappa(t, X_2) \\ &\kappa(t, X_n) \uparrow \kappa(t, X) \qquad \text{if } 0 \le X_1 \le X_2 \le \ldots \uparrow X \end{aligned}$$

  Also require functions of T to act like constants?
- Only negligible sets stand between $\kappa(t, \cdot)$ and countable additivity.

- Folklore: likelihood ratios are sufficient
- Suppose $p_i(\omega) = d\mathbb{P}_i/d\lambda$ for $i = 1, 2, \ldots, k$
  Define $T : \Omega \to \mathbb{R}_+^k$ by $T(\omega) = \big(p_1(\omega), \ldots, p_k(\omega)\big)$.
- Without loss of generality, λ is a probability measure.
- Let $\kappa(t, X)$ be a version of $\mathbb{E}(X \mid T = t)$ for expectations under λ. Sufficient?

- Can $\kappa(t, \cdot)$ be represented by a (finitely additive? countably additive?) probability π_t?

- cf. Halmos (1950, problem 48.4)
  Let λ = Lebesgue on Borel sigma-field $\mathcal{B}$ of $[0, 1]$.
  Fix a set A with $\lambda^* A = 1$ and $\lambda_* A = 0$. Extend λ to a probability on the sigma-field $\mathcal{F}$ generated by $\mathcal{B}$ and A:

$$\lambda\left(AB_1 + A^c B_2\right) = \tfrac{1}{2}\left(\lambda(B_1) + \lambda(B_2)\right) \qquad \text{for all } B_i \in \mathcal{B}.$$

  Take $p_1(\omega) = 2\omega$ and $p_2(\omega) = 2(1 - \omega)$.

- Counterexample contrived? Violates typical regularity properties? Why not always assume enough regularity properties to eliminate need for Kolmogorov conditional expectations?

- Advantages of abstraction:
  Simplifications if we are only interested in a small collection of rvs X.
  Reduces to Markov kernel under extra regularity conditions.
  Domination and separability assumptions give at least a finitely additive π_t
  Composition can lead to Markov kernels.
- Disadvantages:
  Can $\kappa(t, \cdot)$ really be regarded as a randomizing mechanism?
  Need topological assumptions to make π_t countably additive.
  Difficulties when Θ not finite or model not dominated.

- Countable additivity depends on choice of Ω.
- Why not Lebesgue measure on $\Omega = \{r_1, r_2, \ldots\}$ = rationals in $(0, 1]$?
  Define $\lambda(r, s] = s - r$. Extend by finite additivity.
  Unfortunately $A_n := \Omega \backslash \{r_1, r_2, \ldots, r_n\} \downarrow \emptyset$ but $\lambda A_n = 1$ for all n.
- Rescue countable additivity by adding more points to Ω so that $\cap_n A_n$ no longer empty. Add enough new points to "neutralize" decreasing sequences that would violate countable additivity.
- Try to make finitely additive measures on original Ω correspond to countably additive measures on augmented Ω.

- Markov kernels: $\omega \sim P \mapsto \boxed{KP := \int K_\omega(\cdot) P(d\omega)} \mapsto y \sim KP$
- (Le Cam) "Generalized randomizations", $P \mapsto KP$: increasing, linear, preserve total mass (also work when $P\Omega \neq 1$).
- Advantages of Le Cam abstraction:
  Reduces to Kolmogorov under regularity conditions.
  Reduces to Markov kernel under more regularity conditions.
  Domination and separability assumptions plus ... give at least a finitely additive π_t
  Composition can lead to Markov kernels.
  Don't need extra topological assumptions to get nice existence theorems.
  No extra difficulties when Θ not finite or model not dominated.

$\mathcal{P}$ model: probabilities on $[0, 2)$,

$$\mathbb{P}_\theta = \begin{cases} \text{point mass at } 1 + \theta & \text{for } 0 < \theta \leq 1 \\ \text{Lebesgue measure on } [0, 1) & \text{for } \theta = 0 \end{cases}$$

$\mathcal{Q}$ model: probabilities on $[0, 1)$,

$$\mathbb{Q}_\theta = \begin{cases} \text{point mass at } \theta & \text{for } 0 < \theta \leq 1 \\ \text{Lebesgue measure on } [0, 1) & \text{for } \theta = 0 \end{cases}$$

- How to test hypothesis $\theta = 0$ versus $0 < \theta \leq 1$?
- $\mathcal{A}$ = sigma-field generated by singletons. Sufficient? Pairwise sufficient? (Torgersen 1991, Section 1.5).

Some ways to use Le Cam framework:

- Add regularity assumptions so that all generalized objects reduce to their classial analogs.
- Use Le Cam framework as a convenient way of finding traditional solutions:
  (i) Find generalized solution. (ii) Show that solution from (i) can actually be identified with a traditional solution.
- Rethink what we mean by a statistical model. For example, what does it mean to say that data are observations on a fixed distribution? Why are sample spaces needed? ...

References

Halmos, P. R. (1950), *Measure Theory*, Van Nostrand, New York, NY. July 1969 reprinting.

Le Cam, L. (1986), *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York.

Torgersen, E. (1991), *Comparison of Statistical Experiments*, Cambridge University Press.

van der Vaart, A. (2002), 'The statistical work of Lucien Le Cam', *Annals of Statistics* **30**, 631–682.