where $x = \rho^2[1 + a^2/(N + 1)]^{-1}$ and $I_x(\frac{1}{2}(N + 1), \frac{1}{2})$ is Karl Pearson's notation for the Incomplete Beta-Function as tabled in [2].

In the preceding discussion it has been assumed that the mean of the process (x_t) is known to be zero. If the mean must be estimated from the sample, the serial correlation coefficient will be

$$r' = \frac{\sum_{t=1}^{N} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^{N} (x_t - \bar{x})^2}, \qquad \bar{x} = \frac{\sum_{t=1}^{N} x_t}{N}.$$

All of the results concerning r also hold true for r' with N degrees of freedom rather than $N + 1$.

REFERENCES

(1) R. B. Leipnik, "Distribution of the serial correlation coefficient in a circularly correlated universe," Ann. Math. Stat., vol. 18 (1947).
(2) K. Pearson, *Tables of the Incomplete Beta-Function* Cambridge, 1934.

GROUPS AND CONDITIONAL MONTE CARLO

By J. G. Wendel

University of Michigan

Summary. The conditional Monte Carlo technique advanced by Tukey *et al.* [1, 2] has been explained in analytic terms by Hammersley [3]. This note offers an alternative explanation, wherein the group-theoretic aspect of the problem plays the dominant role. The method is illustrated on an example simpler than that treated in [1, 2].

The framework. Throughout this note X will be a random vector in euclidean n-space $\mathfrak{X}$, having distribution function G. F will denote a distribution function absolutely continuous with respect to G, with Radon-Nikodym derivative $dF/dG \equiv w$, so that

$$F(M) = \int_M w(x)\, dG(x)$$

for all Borel sets M, and

$$\int \varphi(x)\, dF(x) = \int \varphi(x) w(x)\, dG(x)$$

for Borel functions φ. It is standard in this situation to call w a *weight* and to say that X (drawn from G) *with weight* $w(X)$ *is a sample from* F; thus for Borel φ we have

$$E_G(\varphi(X)w(X)) = E_F(\varphi(X))$$

Received January 28, 1957; revised April 12, 1957.

where the subscripts on the expectation operators specify the population from which X is sampled. From now on we assume without essential loss of generality that F and G come from densities f and g, so that $w(x) = f(x)/g(x)$ or 0 according as $g(x) \neq 0$ or $= 0$.

The problem. A more interesting case arises when *conditional* expectations are desired. It is by no means apparent that it will in general be possible to find a weight-function w^* and an appropriate modification X^* of X so that

$$E_g(\varphi(X^*)w^*(X)) = E_f(\varphi(X) \mid \text{condition on } X)$$

identically in φ, but in fact the main theme of [1, 2] was that highly non-trivial instances do exist. Their problem can be put as follows:

$\mathfrak{X}$: Euclidean n-space;

$\mathfrak{A}$: a locally compact non-necessarily Abelian group of 1-1 differentiable transformations acting on $\mathfrak{X}$, such that the mapping $(\alpha, x) \rightarrow \alpha x (\alpha \in \mathfrak{A}, x \in \mathfrak{X})$ is measurable;

$\partial(\alpha x)/\partial x$: the Jacobian of $\alpha \in \mathfrak{A}$ at $x \in \mathfrak{X}$;

$dm(\alpha)$: a fixed right-invariant Haar measure over $\mathfrak{A}$;

A: a left-homogeneous mapping defined on almost all of $\mathfrak{X}$ onto $\mathfrak{A}$, so that $A(\alpha x) = \alpha A(x)$ for all α and all x in the domain of A;

σ: the density function of A, assumed to exist; thus, for Borel sets $B \subset \mathfrak{A}$

$$\Pr \{A(X) \in B\} = \int_B \sigma(\alpha)\, dm(\alpha) = \int I_B(A(x))f(x)\, dx$$

where I_B is the indicator of B, and X of course has density f.

In [1, 2] the group $\mathfrak{A}$ consisted of the multiplicative group of positive reals, acting as dilations on $\mathfrak{X}$; then $dm(\alpha)$ can be taken to be $d\alpha/\alpha$, and the Jacobian is just α^n.

The problem is to express $E_f(\varphi(X) \mid A(X) = \alpha_0)$ as an unconditional expectation $E_g(\varphi(X^*)w^*(X))$, where X is sampled from density g, X^* is a suitable modification of X, and w^* is an appropriate weight. This is certainly natural in the Monte Carlo setting, for it would save us from having to waste most of our observations, namely those X for which $A(X)$ is not reasonably close to α_0.

Development of solution. (The formulae set down in this section are those of [1], but interpreted in the broader setting and subjected to formal proof.)

In view of the homogeneity of A the obvious choice of modification $X \rightarrow X^*$ is to force the condition $A(X^*) = \alpha_0$ to hold. This will be achieved if we take $X^* = \alpha_0 A(X)^{-1} X = \alpha X$, where α denotes $\alpha_0 A(X)^{-1}$ and is, like X, a random variable. Finding the appropriate weights will occupy the remainder of this section.

LEMMA. *Suppose that X with weight $w(X)$ is a sample from density f. Then for each $\alpha \in \mathfrak{A}$, αX with weight $w(\alpha, X)$ is a sample from f too, where*

$$w(\alpha, x) = w(x)\{f(\alpha x)/f(x)\} \mid \partial(\alpha x)/\partial x \mid.$$

PROOF. Write $Y = \alpha X$. We want to evaluate $\int \varphi(y)f(y)\,dy$ for Borel φ. But this is

$$\begin{aligned}\int \varphi(\alpha x)f(\alpha x)\,d(\alpha x) &= \int \varphi(\alpha x)f(\alpha x) \mid \partial(\alpha x)/\partial x \mid dx \\ &= \int \varphi(\alpha x)\{f(x)/g(x)\}\{f(\alpha x)/f(x)\}\mid \partial(\alpha x)/\partial x \mid g(x)\,dx \\ &= \int \varphi(\alpha x)w(x)\{f(\alpha x)/f(x)\}\mid \partial(\alpha x)/\partial x \mid g(x)\,dx \\ &= \int \varphi(\alpha x)w(\alpha, x)g(x)\,dx\end{aligned}$$

as was to be shown.

THEOREM. *Let μ be an arbitrary density over $\mathfrak{A}$. For each $\alpha \in \mathfrak{A}$ suppose that αX with weight $w(\alpha, X)$ is a sample from density f. Then, with $A(x)$ as above, $\alpha = \alpha_0 A(x)^{-1}$ (so that $A(\alpha x) = \alpha_0$) and $X^* = \alpha X$, we have X^* with weight w^* is a sample from f* conditioned *by $A(X^*) = \alpha_0$; the weight w^* is given by*

$$w^*(x) = w^*(\alpha_0, x) = \sigma(\alpha_0)^{-1}\mu(\alpha_0 A(x)^{-1})w(\alpha_0 A(x)^{-1}, x).$$

PROOF. For simplicity write Y in place of X^*. We want to evaluate $E_f(\varphi(Y) \mid A(Y) = \alpha_0) =_{\text{Def}} \psi(\alpha_0)$. ψ is characterized (up to almost-everywhere equivalence) by

$$(1)\qquad \int_B \psi(\alpha)\sigma(\alpha)\,dm(\alpha) = \int_{\{A(y)\in B\}} \varphi(y)f(y)\,dy = \int \tilde{\varphi}(y)f(y)\,dy$$

where $\tilde{\varphi} = \varphi I_{A^{-1}B}$, B a Borel subset of $\mathfrak{A}$. By hypothesis we have for all Borel $\tilde{\varphi}$ and each $\beta \in \mathfrak{A}$,

$$(2)\qquad \int \tilde{\varphi}(y)f(y)\,dy = \int \tilde{\varphi}(\beta x)w(\beta, x)g(x)\,dx.$$

Multiplying both sides of (2) by $\mu(\beta)$ and integrating over $\mathfrak{A}$ with respect to $dm(\beta)$ gives

$$(3)\qquad \begin{aligned}\int \tilde{\varphi}(y)f(y)\,dy &= \int_{\mathfrak{A}} \mu(\beta)\,dm(\beta) \int_{\mathfrak{X}} \tilde{\varphi}(\beta x)w(\beta, x)g(x)\,dx \\ &= \int_{\mathfrak{X}}\int_{\mathfrak{A}} \mu(\beta)\tilde{\varphi}(\beta x)w(\beta, x)\,dm(\beta)g(x)\,dx\end{aligned}$$

where the interchange of order of integration is justified by Fubini's theorem. Putting $\beta = \alpha A(x)^{-1}$ and invoking right invariance replaces the right side by (3) by

$$\int_{\mathfrak{X}}\int_{\mathfrak{A}} \mu(\alpha A(x)^{-1})\tilde{\varphi}(\alpha A(x)^{-1}x)w(\alpha A(x)^{-1}, x)\,dm(\alpha)g(x)\,dx$$

$$= \int_{\mathfrak{X}} \int_B \mu(\alpha A(x)^{-1})\varphi(\alpha A(x)^{-1}x)w(\alpha A(x)^{-1}, x)\, dm(\alpha)g(x)\, dx$$

$$= \int_B \sigma(\alpha)\, dm(\alpha) \int_{\mathfrak{X}} \sigma(\alpha)^{-1}\mu(\alpha A(x)^{-1})\varphi(\alpha A(x)^{-1}x)w(\alpha A(x)^{-1}, x)g(x)\, dx$$

$$= \int_B \sigma(\alpha)\, dm(\alpha) \int_{\mathfrak{X}} w^*(x)\varphi(\alpha A(x)^{-1}x)g(x)\, dx$$

and the result now follows on comparison with the left member of (1).

The solution. Combining the formulae of lemma and theorem shows that, for X drawn from G and weighted by

$$\begin{aligned} w^*(x) &= w^*(\alpha_0, x) \\ &= \sigma(\alpha_0)^{-1}w(x)\{f(\alpha_0 A(x)^{-1}x)/f(x)\} \mid \partial(\alpha_0 A(x)^{-1}x)/\partial x \mid \mu(\alpha_0 A(x)^{-1}), \end{aligned}$$

the average of $\varphi(X^*)w^*(X)$ yields the desired conditional expectation, where $X^* = \alpha X$, with α chosen so that $A(X^*) = A(\alpha X) = \alpha_0$. The arbitrariness of μ may seem peculiar, but its role may be clarified by the following simple example.

Example. Let X be a scalar random variable with density f. It is desired to find $E_f(\varphi(X) \mid X > 0)$ by sampling. To put this in the present framework let $\mathfrak{A}$ = two-element group of numbers $+1$, -1 acting multiplicatively on $\mathfrak{X}$, the reals. Let $A(x) = +$ or -1 according as $x >$ or < 0, with $A(0)$ arbitrary. Then the homogeneity property is clearly satisfied, and the problem amounts to finding $E_f(\varphi(X) \mid A(X) = +1)$. The Haar measure $dm(\alpha)$ over $\mathfrak{A}$ is defined by placing unit mass at each of $+1$, -1. The density σ is then

$$\sigma(+1) = p = \Pr\{A(X) = +1\} = \Pr\{X > 0\} = \int_0^\infty f(x)\, dx,$$

$$\sigma(-1) = 1 - p.$$

For $\mu(\alpha)$ we pick a number λ between zero and one inclusive and set $\mu(+1) = \lambda$, $\mu(-1) = 1 - \lambda$. The weight $w(x)$ is identically one, as is the absolute value of the Jacobian. Substituting all this information into the formula for w^* we obtain

$$w^*(+1, x) = w^*(x) = (1/p)\{f(A(x)^{-1}x)/f(x)\}\mu(A(x)^{-1}x),$$

i.e.
$$w^*(x) = \begin{cases} \lambda/p & \text{if } x > 0, \\ (1/p)\{f(-x)/f(x)\}(1 - \lambda) & \text{if } x < 0. \end{cases}$$

It is easy to verify directly now that $E_f(\varphi(X^*)w^*(X)) = E_f(\varphi(X) \mid X > 0)$; the Monte Carlo procedure will be: observe values $X = x_1, x_2, \cdots, x_n$ from the density f; for each x_i compute $x_i^* = \mid x_i \mid$ and $w^*(x_i)$ the appropriate expression from above, and use $(1/n)\sum\varphi(\mid x_i \mid)w^*(x_i)$ as the desired estimate.

Naturally one would wish to choose λ so as to minimize the variance or—what comes to the same thing in view of unbiassedness—some multiple of the mean square of the estimator. We find that

$$p^2E(\{\varphi w^*\}^2) = \lambda^2 \int_0^\infty \varphi(y)^2 f(y)\, dy + (1 - \lambda)^2 \int_0^\infty \frac{f(y)}{f(-y)} \varphi(y)^2 f(y)\, dy$$
$$= \lambda^2 J_1 + (1 - \lambda)^2 J_2,$$

say, which is minimized by setting

$$\lambda = J_2/(J_1 + J_2).$$

In case f is symmetric $J_1 = J_2$ and optimum $\lambda = 1/2$; here the naive procedure of rejecting negative x's corresponds to $\lambda = 1$ and maximizes the mean square! However, if $f(-y) = 0$ over a stretch in which $f(y) > 0$ then $J_2 = \infty$, and we must take $\lambda = 1$, adopt the naive solution, in order to obtain finite variance of estimate. Finally, in case $\varphi(y)$ and $f(-y)/f(y)$ have large similar peaks near some $y_0 > 0$ then J_1 may be very much larger than J_2 and optimum λ very close to 0.

REFERENCES

[1] H. F. Trotter and J. W. Tukey, "Conditional Monte Carlo for normal samples." *Symposium on Monte Carlo Methods*, New York, 1956, pp 64–79.
[2] H. J. Arnold, B. D. Bucher, H. F. Trotter, and J. W. Tukey, "Monte Carlo techniques in a complex problem about normal samples," *Symposium on Monte Carlo Methods*, New York, 1956, pp. 80–88.
[3] J. M. Hammersley, "Conditional Monte Carlo," *J. Assoc. Comp. Mach.* Vol. 3 (1956), pp. 73–76.

TABLES FOR TYPE A CRITICAL REGIONS

By Harry Weingarten

Bureau of Ships

1. This note provides tables connected with work by Neyman [1] and Johnson [2] on testing hypotheses, expanding the table given in [2]. This table, as expanded provides solutions for the values of A satisfying,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{A-Bu^2}^{\infty} e^{-\frac{1}{2}u^2 - \frac{1}{2}v^2}\, dv\, du = \alpha$$

for $\alpha = .01, .05$, and $B = 0(.1)5, 5(1)10, 10(10)100$.
When $\alpha = .05$ set $A = 3.8414588B + \rho_{.05}$, and when
$\alpha = .01$ set $A = 6.6348966B + \rho_{.01}$

Received March 14, 1957.