A Curious Connection Between Optimal Stopping and Branching Processes

David Assaf, Larry Goldstein and Ester Samuel-Cahn

Abstract

A curious connection exists between the theory of branching processes and optimal stopping for independent random variables. For the branching process Z_n with offspring distribution Y, there exists a random variable X such that the probability $P(Z_n = 0)$ of extinction of the n^{th} generation in the branching process equals the value obtained by optimally stopping the independent sequence $X_1, \ldots, X_n$ with distribution X. The correspondence can be generalized to the inhomogeneous and infinite horizon cases, and sometimes furnishes simple 'stopping rule' methods for computing various characteristics of branching processes, such as rates of convergence of the n^{th} generation's extinction probability to the eventual extinction probability. It also provides a method for computing bounds on branching extinction probabilities by applying prophet inequalities and suboptimal rules in the corresponding stopping problem. This curious correspondence may also be used in the other direction, to inform the theory of optimal stopping using results from branching. We have no probabilistic explanation for the correspondence.

Two Choice Stopping: How much is an extra chance worth?

David Assaf, Larry Goldstein and Ester Samuel-Cahn

Abstract

A statistician sequentially observes the values in the independent, identically distributed sequence $X_n, \ldots, X_1$ with known distribution F, and is given two chances to choose as small a value as possible using stopping rules. Let V_n^2 equal the expectation of the smaller of the two values chosen when proceeding optimally. Then for a large class of F's belonging to the domain of attraction (for the minimum) of $\mathcal{D}(G^\alpha)$, where $G^\alpha(x) = [1 - \exp(-x^\alpha)]\mathbf{I}(x \geq 0)$,

$$\lim_{n\to\infty} nF(V_n^2) = h^\alpha(d_\alpha)$$

where $d_\alpha > 0$ is the unique solution d to

$$\int_0^d h(y)dy + (1/\alpha - d)h(d) = 0,$$

and $h(y)$ is the function

$$h(y) = \left(\frac{y}{1 + \alpha y/(\alpha+1)}\right)^{1/\alpha} \quad \text{for } y \geq 0.$$

For 'most' α, having two choices is a substantial improvment over having one as measured by asymptotic distance to the "prophet" sequence $E(\min\{X_n, \ldots, X_1\})$.

Berry Esseen Bounds for Local Extremes and Combinatorial Central Limit Theorems, using Size and Zero Biasing

Larry Goldstein

Abstract

For $W \geq 0$ with finite mean μ, say W^s has the W size biased distribution if

$$EWf(W) = \mu Ef(W^s) \quad \text{for all } f,$$

and similarly, for W with mean zero and finite positive variance σ^2, say W^* has the W zero biased distribution if

$$EWf(W) = \sigma^2 Ef'(W^*) \quad \text{for all } f.$$

By Stein's method a non-negative, or mean zero, W is close to normal when it can be coupled closely to its size bias, or zero bias, version respectively. By the use of smoothing inequalities, Berry Esseen bounds are developed for both types of couplings, and are illustrated for the normal approximation of the number of local extreme values on a graph, and for a combinatorial central limit theorem with random permutation having distribution constant on cycle type.

Normal Approximation for Hierarchical Sequences

Larry Goldstein

Abstract

Given $F : [a,b]^k \to [a,b]$ and a non-constant X_0 with $P(X_0 \in [a,b]) = 1$, define the hierarchical sequence of random variables $\{X_n\}_{n\geq 0}$ by $X_{n+1} = F(X_{n,1}, \ldots, X_{n,k})$, where $X_{n,i}$ are i.i.d. as X_n. Such sequences arise from self similar structures which have been extensively studied in the physics literature to model, for example, the conductivity of a random medium. Under an averaging and smoothness condition on non-trivial F, a upper bound of the form $C\gamma^n$ for $0 < \gamma < 1$, is obtained on the Wasserstein distance between the standardized distribution of X_n and the normal. The results apply, for instance, to random resistor networks, and introducing the notion of strict averaging, to hierarchical sequences generated by certain compositions. As an illustration, upper bounds on the rate of convergence to the normal are derived for the hierarchical sequence generated by the weighted diamond lattice, which is shown to exhibit a full range of convergence rate behavior.

Sampling Case-Control Studies Over Time

Ornulf Borgan, Larry Goldstein, Bryan Langholz and Haimeng Zhang

Abstract

Sampling is inevitable in epidemiological studies which follow large cohorts over time in order to detect potential relationships between exposure and disease. Variety in sampling methods can range from the Nested Case-Control design, consisting of taking a simple random sample from those at risk to serve as controls for any failure, to the Counter Matching design, where the sampling results in a fixed number from each cohort strata. As designs such as Counter Matching use to great advantage any easily available cohort information which might serve as a surrogate for true exposure, it is valuable to have tools which provide for a general analysis of studies under a wide range of sampling methods. Adopting the Cox proportional hazard model on the cohort, such general methods can be provided for regression parameter and cumulative baseline hazard estimation when the cohort is sampled according to a predictable sampling probability law. The key to the methodology is to define counting processes which count joint failure and sampled risk sets occurrences and to derive the appropriate intensities for these counting processes, leading to analysis by martingale methods parallel to those for full cohort data. Asymptotic relative efficiencies for the counter matching design relative to the full cohort demonstrate that significant efficiency gains can be made by using available surrogate information. The designs and their counting process analysis carry over directly to the Mantel Haenszel estimator, generalizing that framework not only to sampling but also to models with any finite number of exposure levels. Questions of absolute, that is, semi-parametric efficiency, are touched on.

Frequency Matching in Logistic Regression, Rejective Sampling, and Local Central Limit Theorems

Richard Arratia, Larry Goldstein and Bryan Langholz

Abstract

Designs such as Nested Case-Control sampling and Counter Matching were originally developed for use in semi-parametric models where individuals with a common base line hazard are followed over time in order to detect relationships between exposure and disease. Though such designs can be implemented in studies having no time dependence, martingale methods and counting process techniques can no longer be applied. In their place, local central limit theorems for independent but not identically distributed Bernoulli random variables must be developed, which by the appropriate conditioning, allows for the derivation of asymptotic estimator properties, and in addition provides results about the high correlation structure in the sampling proportional to size type scheme known as rejective sampling. The resulting properties of the estimators so obtained, and relative efficiencies of designs such as Counter Matching in particular, show as in the continuous time case that substantial efficiency gains may be achieved by the use of surrogate information.