SELF-COUPLINGS AND THE CONCENTRATION FUNCTION

Svante Janson

A coupling of a random variable X with itself is a pair of random variables (X', X''), defined on a common probability space, such that $X' \overset{\mathrm{d}}{=} X'' \overset{\mathrm{d}}{=} X$. There are two obvious such couplings, viz. the trivial with $X' = X'' = X$ and the one consisting of two independent copies of X. It is less obvious that other couplings of X with itself may be of interest, but the following theorem shows that they may be used to derive estimates for the concentration function $Q_X(h)$, which we recall is defined as the left hand side of (1).

Theorem. *For any $h \geq 0$,*

$$\sup_{-\infty<a<\infty} \mathrm{P}(a \leq X \leq a+h) = \inf_{X' \overset{\mathrm{d}}{=} X'' \overset{\mathrm{d}}{=} X} \mathrm{P}(X'' \leq X' + h). \tag{1}$$

Proof. Denote the left and right hand sides of (1) by l and r, respectively.

If $X'' > X' + h$, then $X' < a$ or $X'' > X' + h \geq a + h$. Thus, for any a,

$$\mathrm{P}(X'' > X' + h) \leq \mathrm{P}(X' < a) + \mathrm{P}(X'' > a + h) = 1 - \mathrm{P}(a \leq X \leq a + h),$$

which proves that $l \leq r$.

For the opposite inequality, let $l < s < 1$ (the case $l = 1$ is trivial), let U be a random variable that is uniformly distributed on $(0,1)$ and let F be the distribution function of X and F^{-1} its right-continuous inverse. Define $X' = F^{-1}(U)$ and

$$X'' = \begin{cases} F^{-1}(U+s), & U < 1-s; \\ F^{-1}(U+s-1), & U > 1-s. \end{cases}$$

It is easily verified that $X' \overset{\mathrm{d}}{=} X'' \overset{\mathrm{d}}{=} X$ and that $\mathrm{P}(X'' \leq X' + h) = \mathrm{P}(1 - s < U < 1) = s$. Hence $r \leq s$, which implies $r \leq l$. The proof is complete.

Remark. We similarly have

$$\sup_{-\infty<a<\infty} \mathrm{P}(a < X < a+h) = \sup_{-\infty<a<\infty} \mathrm{P}(a < X \leq a+h) = \inf_{X' \overset{\mathrm{d}}{=} X'' \overset{\mathrm{d}}{=} X} \mathrm{P}(X'' < X' + h). \tag{2}$$

The supremum in (1) and the infimum in (2) are attained, and may thus be written as max and min, while simple examples show that this is not always true for the infimum in (1) and the suprema in (2).

An application of this theorem is given in [1], which studies the concentration of the length of the longest increasing subsequence in a random permutation of a given length n. Two couplings (X', Y') and (X'', Y'') are constructed between this variable X and another random variable Y (actually the same variable for a different n) such that for certain $h = h(n)$ and $\varepsilon = \varepsilon(n)$, $\mathrm{P}(Y' > X' + h) \geq 1 - \varepsilon$ and $\mathrm{P}(X'' \neq Y'') < \varepsilon$. We may assume that X', X'', Y', Y'' are defined on the same probability space with $Y' = Y''$; then (X', X'') is a coupling of X with itself such that $P(X'' \leq X' + h) < 2\varepsilon$. The theorem gives the estimate $Q_X(h) < 2\varepsilon$.

Supported by the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine

Reference

1. B. Bollobás and S. Janson, On the length of the longest increasing subsequence in a random permutation, in preparation.