THE SECOND MOMENT METHOD, CONDITIONING AND APPROXIMATION

Svante Janson

Abstract. The second moment method has been fruitfully combined with conditioning by several authors, for example Robinson and Wormald [7], [8] and Frieze and Janson [2]. We try to give a general exposition of this method, and of its relation to approximations of random variables by simpler variables. We discuss several examples where the method has been used in different forms.

1. The second moment method

One of the classical methods to show the existence (with large probability) of certain substructures in a random graph is the second moment method. We can describe a general version of the method as follows.

Consider some random combinatorial structure G. In order to study the existence of a certain type of substructure (for example, a Hamilton cycle, a perfect matching, an induced $C_4, \dots$), we let $X = X(G)$ be the number of such substructures in G. Thus X is a random variable taking non-negative integer values, and the problem is to estimate $\mathrm{P}(X > 0)$. We assume that $\mathrm{P}(X > 0) > 0$ (otherwise the substructures never appear) and that $\mathrm{E}\,X^2 < \infty$ (in fact, we usually study finite random structures, where X is bounded).

Typically, we are interested in asymptotic results. We consider a sequence of such random structures G (a more precise notation would be G_n, but we will usually ignore the subscript n) and want to show $\mathrm{P}(X > 0) \to 1$. The second moment method asserts that this follows if we can show

$$\operatorname{Var} X/(\mathrm{E}\,X)^2 \to 0 \tag{1}$$

or, equivalently,

$$\mathrm{E}\,X^2/(\mathrm{E}\,X)^2 \to 1. \tag{1'}$$

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

Research carried out during a visit to the IMA, Minneapolis

The proof is immediate by Chebyshev's inequality:

$$\mathrm{P}(X=0) \le \frac{\operatorname{Var} X}{(\mathrm{E}\, X)^2} = \frac{\mathrm{E}\, X^2}{(\mathrm{E}\, X)^2} - 1. \tag{2}$$

Alternatively, the Cauchy–Schwarz inequality gives

$$(\mathrm{E}\, X)^2 = \bigl(\mathrm{E}(X I(X>0))\bigr)^2 \le \mathrm{E}\, X^2\, \mathrm{P}(X>0)$$

and thus

$$\mathrm{P}(X=0) = 1 - \mathrm{P}(X>0) \le 1 - \frac{(\mathrm{E}\, X)^2}{\mathrm{E}\, X^2}. \tag{3}$$

The bound in (3) is better than the bound in (2), but for our purposes both bounds are equally powerful.

Note that (1) is much stronger than the conclusion $\mathrm{P}(X>0) \to 1$. In fact, (1) can be written $\operatorname{Var}(X/\mathrm{E}\, X) \to 0$ or $\mathrm{E}(X/\mathrm{E}\, X - 1)^2 \to 0$, i.e.

$$X/\mathrm{E}\, X \xrightarrow{L^2} 1. \tag{4}$$

In particular, by Chebyshev's inequality again,

$$X/\mathrm{E}\, X \xrightarrow{p} 1; \tag{5}$$

i.e., for every $\varepsilon > 0$,

$$\mathrm{P}\bigl((1-\varepsilon)\,\mathrm{E}\, X < X < (1+\varepsilon)\,\mathrm{E}\, X\bigr) \to 1. \tag{6}$$

Hence, the second moment method is applicable only when X can be well approximated (in the sense of (4)–(6)) by the non-random value $\mathrm{E}\, X$; in other words, the distribution of X has to be (asymptotically) concentrated at a single value.

2. Conditioning

There are many cases where the second moment method fails because (4)–(6) do not hold; for example, $X/\mathrm{E}\, X$ may converge in distribution to a non-degenerate random variable. In such cases, the method may be combined with conditioning as follows.

Let us choose another random variable $Y = Y(G)$; Y does not have to be real-valued, for example, Y may be a sequence $(Y^{(1)}, Y^{(2)}, \dots)$ of real random variables. For simplicity we assume that Y is discrete, i.e. Y assumes only a countable number of values. (For the approach to be useful, we also want Y to be simpler than X in some sense.) Two examples, considered again later, are $Y =$ the number of edges in a random graph $G(n,p)$, and $Y = (d_1, \dots, d_n)$, where d_i is the degree of vertex i in a random graph or hypergraph.

Recall that the conditional expectation $\mathrm{E}(X \mid Y)$ is a random variable which is a function of Y; more precisely, it equals $g(Y)$, where the function g is defined by $g(y) = \mathrm{E}(X \mid Y = y)$ whenever $\mathrm{P}(Y = y) > 0$. Similarly, the conditional variance is the random variable

$$\mathrm{Var}(X \mid Y) = \mathrm{E}\big((X - \mathrm{E}(X \mid Y))^2 \mid Y\big) = \mathrm{E}(X^2 \mid Y) - \mathrm{E}(X \mid Y)^2.$$

We can now apply (2) after conditioning on Y (i.e., conditioning on $Y = y$ for each y) and obtain

$$\mathrm{P}(X = 0 \mid Y) \leq \frac{\mathrm{Var}(X \mid Y)}{\mathrm{E}(X \mid Y)^2} \tag{7}$$

whenever $\mathrm{E}(X \mid Y) > 0$. (Recall that both sides of (7) are random variables.)

For convenience, we make the right hand side of (7) everywhere defined by assigning an arbitrary value in $[0, \infty]$ to it when $\mathrm{E}(X \mid Y) = 0$ (in which case it is of the form $0/0$).

Theorem 1. *Suppose that* $\mathrm{P}\big(\mathrm{E}(X \mid Y) = 0\big) \to 0$ *and*

$$\mathrm{Var}(X \mid Y) / \mathrm{E}(X \mid Y)^2 \xrightarrow{p} 0. \tag{8}$$

Then $\mathrm{P}(X = 0) \to 0$.

Proof. Since conditional probabilities are bounded by 1, (7) can be rewritten

$$\mathrm{P}(X = 0 \mid Y) \leq \min\big(1, \mathrm{Var}(X \mid Y) / \mathrm{E}(X \mid Y)^2\big) + I\big(\mathrm{E}(X \mid Y) = 0\big) \tag{9}$$

which by taking expectations yields

$$\mathrm{P}(X = 0) \leq \mathrm{E}\min\big(1, \mathrm{Var}(X \mid Y) / \mathrm{E}(X \mid Y)^2\big) + \mathrm{P}\big(\mathrm{E}(X \mid Y) = 0\big). \tag{10}$$

The first term on the right hand side tends to 0 by (8) and dominated convergence, while the second tends to 0 by assumption. $\square$

The condition (8) may be rewritten

$$\mathrm{E}(X^2 \mid Y) / \mathrm{E}(X \mid Y)^2 \xrightarrow{p} 1 \tag{11}$$

or

$$\mathrm{P}\big(\mathrm{Var}(X \mid Y) > \varepsilon\, \mathrm{E}(X \mid Y)^2\big) \to 0, \qquad \text{for every } \varepsilon > 0. \tag{12}$$

We may also write it as

$$\mathrm{Var}\Big(\frac{X}{\mathrm{E}(X \mid Y)} \mid Y\Big) \xrightarrow{p} 0 \tag{13}$$

and it follows easily that the conditions of Theorem 1 imply

$$\frac{X}{\mathrm{E}(X \mid Y)} \xrightarrow{p} 1 \tag{14}$$

or, for every $\varepsilon > 0$,

$$\mathrm{P}\big((1-\varepsilon)\,\mathrm{E}(X \mid Y) < X < (1+\varepsilon)\,\mathrm{E}(X \mid Y)\big) \to 1. \tag{15}$$

Hence the conditional second moment method is applicable only when X can be well appproximated (with a relative error that is small in probability) by a function of Y. The converse is not necessarily true, since variables may converge in probability without converging in L^2, but we do not expect this to happen in practical applications, and thus we expect Theorem 1 to be applicable whenever X is well approximated by some function $g(Y)$. (We may then take $g(Y) = \mathrm{E}(X \mid Y)$, but some approximation of $\mathrm{E}(X \mid Y)$ may be simpler.) This should be compared to the unconditional second moment method described above, which applies only when the approximation can be chosen as a "constant". (We put "constant" in inverted commas, since we really consider a sequence $G = G_n$ of random structures and everything, including the function g and this "constant", may depend on the parameter n.)

We also remark that when the conditional second moment method is applicable, it typically also yields the asymptotic distribution of X exactly when the unconditional second moment method is *not* applicable. In fact, we have just seen that then $X/g(Y) \xrightarrow{p} 1$ for some variable $g(Y)$ which is not (asymptotically) degenerate. Hence, if for example $g(Y)/\,\mathrm{E}\,g(Y)$ or $a_n \log g(Y) - b_n$ (for some constants a_n, b_n) converges to some (non-degenerate) distribution, we obtain the same result for X. (This is closely related to the projection methods for obtaining asymptotic distributions, where, however, one often considers approximations of a variable X by a *linear* function $g(Y)$ only.)

3. Applications

We have so far discussed the theoretical limits of the method. For practical applications, we also have to be able to verify the hypotheses of the theorem. There are at least three different methods to do this. We describe these methods below, and illustrate them by examples where they have been used by various authors (at least implicitly; the descriptions below are reformulations of the arguments in the original references). The methods often lead to long and non-trivial combinatorial calculations, which we omit here; see the references given in the examples below.

I. The first possibility is a straightforward computation of $\mathrm{E}(X \mid Y)$ and $\mathrm{E}(X^2 \mid Y)$.

Example 1. Janson [3] considered the number of spanning subtrees of the random graph $G(n,p)$, with p constant and $n \to \infty$; let us denote this number by X. It is well-known that $\mathrm{P}(X > 0) \to 1$, but the objective in [3] was to find the asymptotic distribution. This was achieved using the argument above, with Y being the number of edges in $G(n,p)$. Note that conditioning on $Y = m$ is the same as considering the random graph $G(n,m)$; hence $\mathrm{E}(X \mid Y = m)$ and $\mathrm{E}(X^2 \mid Y = m)$ can be evaluated as the first two moments of the number of spanning trees in $G(n,m)$. These can be estimated rather easily, see Moon [6], which leads to an asymptotic log-normal distribution of X.

Example 2. Consider the random hypergraph $G(ns, m; s)$ with ns vertices and m random hyperedges consisting of s vertices each. Frieze and Janson [2] showed that if $s \geq 2$ is fixed, and $n \to \infty$ with $m/n^{4/3} \to \infty$, then the random hypergraph contains a perfect matching (i.e., a covering of the vertex set by n disjoint hyperedges) with probability tending to 1. The proof is based on Theorem 1, letting X be the number of perfect matchings and $Y = (d_1, \dots, d_{ns})$, where d_i is the degree of vertex i, i.e., the number of hyperedges containing i. Actually, the computations are performed for a slightly different model, namely a random multihypergraph where the hyperedges are obtained by sampling ms vertices with replacement and then grouping them s by s; this allows both multiple hyperedges and defect hyperedges with less than s distinct vertices, but the distribution of Y (which becomes multinomial) and the calculation of $\mathrm{E}(X \mid Y)$ and $\mathrm{E}(X^2 \mid Y)$ are much simplified.

It turns out that if $m \gg n^{3/2}$, then $\mathrm{E}(X \mid Y)$ is concentrated at $\mathrm{E}\,X$, and the unconditional second moment method applies, see Schmidt and Shamir [9]. When m is smaller, this is no longer true and we need the conditional version. Note also that the proof requires $m \gg n^{4/3}$, although there is no reason to believe that this is necessary for the result (presumably $m \gg n \log n$ is enough, as it is for $s = 2$). The discussion above gives (or at least suggests) the explanation that the number of perfect matchings is well approximated by a function of the vertex degrees only when $m \gg n^{4/3}$ (and by a constant when $m \gg n^{3/2}$); for smaller m the number depends significantly also on other properties of the hypergraph, but it may of course still be non-zero.

In principle it seems possible to improve the result by conditioning on more parameters beside the vertex degrees, cf. Example 4 below, but it seems hopeless to perform the necessary computations.

II. It is often complicated to calculate the conditional variance $\mathrm{Var}(X \mid Y)$ or $\mathrm{E}(X^2 \mid Y)$. A possible alternative is to use an analysis of variance argument as follows, which only requires calculation of the conditional expectation $\mathrm{E}(X \mid Y)$ and the unconditional variance $\mathrm{Var}\,X$. This is essentially the method invented by Robinson and Wormald [7], [8], see also Cooper, Frieze, Molloy and Reed [1] and Janson [5].

By definition, the terms in the decomposition

$$X = \bigl(X - \mathrm{E}(X \mid Y)\bigr) + \mathrm{E}(X \mid Y) \tag{16}$$

are orthogonal, and

$$\begin{aligned}\operatorname{Var} X &= \operatorname{Var}\bigl(X - \mathrm{E}(X \mid Y)\bigr) + \operatorname{Var}\bigl(\mathrm{E}(X \mid Y)\bigr)\\ &= \mathrm{E}\bigl(X - \mathrm{E}(X \mid Y)\bigr)^2 + \operatorname{Var}\bigl(\mathrm{E}(X \mid Y)\bigr)\\ &= \mathrm{E}\bigl(\operatorname{Var}(X \mid Y)\bigr) + \operatorname{Var}\bigl(\mathrm{E}(X \mid Y)\bigr).\end{aligned} \tag{17}$$

Recall that the model G, and thus also the variables X and Y, depends on a parameter n. The o below should be interpreted as $n \to \infty$.

Theorem 2. *If*

$$\operatorname{Var} X = \operatorname{Var}\bigl(\mathrm{E}(X \mid Y)\bigr) + o\bigl((\mathrm{E}\, X)^2\bigr), \tag{18}$$

and

$$\lim_{\varepsilon\to 0} \limsup_{n\to\infty} \mathrm{P}\bigl(\mathrm{E}(X \mid Y) < \varepsilon\, \mathrm{E}\, X\bigr) = 0, \tag{19}$$

then $\mathrm{P}(X = 0) \to 0$.

Proof. For every $\varepsilon > 0$ and $\delta > 0$,

$$\begin{aligned}&\mathrm{P}\bigl(\operatorname{Var}(X \mid Y)/\,\mathrm{E}(X \mid Y)^2 > \varepsilon\bigr)\\ &\qquad\le \mathrm{P}\bigl(\operatorname{Var}(X \mid Y)/(\mathrm{E}\, X)^2 > \varepsilon\delta^2\bigr) + \mathrm{P}\bigl(\mathrm{E}(X \mid Y) < \delta\, \mathrm{E}\, X\bigr).\end{aligned} \tag{20}$$

Moreover, by (18) and (17), $\mathrm{E}\bigl(\operatorname{Var}(X \mid Y)\bigr) = o(\mathrm{E}\, X^2)$, and thus

$$\mathrm{P}\bigl(\operatorname{Var}(X \mid Y)/(\mathrm{E}\, X)^2 > \varepsilon\delta^2\bigr) \to 0.$$

Hence,

$$\limsup_{n\to\infty} \mathrm{P}\bigl(\operatorname{Var}(X \mid Y)/\,\mathrm{E}(X \mid Y)^2 > \varepsilon\bigr) \le \limsup_{n\to\infty} \mathrm{P}\bigl(\mathrm{E}(X \mid Y) < \delta\, \mathrm{E}\, X\bigr).$$

The right hand side tends to 0 as $\delta \to 0$, and thus the left hand side vanishes for every $\varepsilon > 0$. This shows $\operatorname{Var}(X \mid Y)/\,\mathrm{E}(X \mid Y)^2 \xrightarrow{p} 0$. Furthermore, for every $\varepsilon > 0$,

$$\limsup_{n\to\infty} \mathrm{P}\bigl(\mathrm{E}(X \mid Y) = 0\bigr) \le \limsup_{n\to\infty} \mathrm{P}\bigl(\mathrm{E}(X \mid Y) < \varepsilon\, \mathrm{E}\, X\bigr),$$

and thus (19) implies $\mathrm{P}\bigl(\mathrm{E}(X \mid Y) = 0\bigr) \to 0$. The result now follows by Theorem 1. □

Note that (18) can be regarded as a condition on the variance of the absolute error $X - \mathrm{E}(X \mid Y)$ in the approximation of X by $\mathrm{E}(X \mid Y)$, while condition (8) in Theorem 1 is a condition on the (conditional) variance of the relative error $X/\,\mathrm{E}(X \mid Y) - 1$. The absolute error gives less information when X varies wildly so the theorem is most useful when the variation is not too large. In fact, in all applications so far of this version of the conditional second moment method, we have $\operatorname{Var} X = O\bigl((\mathrm{E}\, X)^2\bigr)$.

Example 3. Robinson and Wormald [7] considered $X =$ the number of Hamilton cycles in a random cubic graph. In principle, they used Theorem 2 with $Y = (\xi_3, \xi_4, \dots)$, where ξ_j is the number of cycles of length j, but they also used a truncation argument. In fact, since $X = \xi_n$ (where n is the order of the graph) is one of the components of Y, $\mathrm{E}(X \mid Y) = X$ so nothing at all is gained by conditioning on Y. Instead, they condition on $Y_M = (\xi_3, \dots, \xi_M)$ and let first $n \to \infty$ for fixed M, and then $M \to \infty$. We can regard their argument as a special case of the following extension of Theorem 2. We assume that Y_M is a variable, defined for some sequence $M \to \infty$ and each G_n.

Theorem 3. *Suppose that*

$$\lim_{M\to\infty} \limsup_{n\to\infty} \bigl[\operatorname{Var} X - \operatorname{Var}\bigl(\mathrm{E}(X \mid Y_M)\bigr)\bigr]/(\mathrm{E}\, X)^2 = 0 \tag{21}$$

and

$$\lim_{\varepsilon\to 0} \limsup_{M\to\infty} \limsup_{n\to\infty} \mathrm{P}\bigl(\mathrm{E}(X \mid Y_M) < \varepsilon\, \mathrm{E}\, X\bigr) = 0. \tag{22}$$

Then $\mathrm{P}(X = 0) \to 0$.

This follows from a corresponding extension of Theorem 1.

Theorem 4. *Suppose that*

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathrm{P}\bigl(\mathrm{E}(X \mid Y_M) = 0\bigr) = 0 \tag{23}$$

and, for each $\varepsilon > 0$,

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathrm{P}\bigl(\operatorname{Var}(X \mid Y_M) > \varepsilon\, \mathrm{E}(X \mid Y_M)^2\bigr) = 0. \tag{24}$$

Then $\mathrm{P}(X = 0) \to 0$.

Proof. By (10),

$$\mathrm{P}(X = 0) \le \mathrm{P}\bigl(\operatorname{Var}(X \mid Y_M) > \varepsilon\, \mathrm{E}(X \mid Y_M)^2\bigr) + \varepsilon + \mathrm{P}\bigl(\mathrm{E}(X \mid Y_M) = 0\bigr).$$

The result follows from (24) and (23) by taking the lim sup as $n \to \infty$, $M \to \infty$ and $\varepsilon \to 0$ (in this order). $\square$

Proof of Theorem 3. Taking the lim sup as $n \to \infty$, $M \to \infty$ and $\delta \to 0$ (in this order) in (20) (with Y_M instead of Y), we obtain (24). Similarly, (23) follows from (22). Hence Theorem 4 applies. $\square$

III. Another way to avoid calculating the conditional variance $\operatorname{Var}(X \mid Y)$ is to estimate the unconditional variance of the variable $X/\mathrm{E}(X \mid Y)$. Note that this variable has mean 1 (if defined to be 1 when $\mathrm{E}(X \mid Y) = 0$). If the variance tends to 0, (13) and (8) follow immediately and we can use Theorem 1 provided $\mathrm{P}\bigl(\mathrm{E}(X \mid Y) = 0\bigr) \to 0$.

The drawback of this version is that we have to calculate (or at least estimate) the variance of the more complicated variable $X/\mathrm{E}(X \mid Y)$. We may sometimes simplify the calculations by using an approximation $g(Y)$ of $\mathrm{E}(X \mid Y)$. (The function g may also depend on the hidden parameter n.) We assume for simplicity that $g(Y) > 0$.

Theorem 5. *Suppose that $g(Y)$ is a positive function of Y such that*

$$\mathrm{E}\,\frac{X}{g(Y)} \to 1, \tag{25}$$

$$\mathrm{Var}\,\frac{X}{g(Y)} \to 0. \tag{26}$$

Then $\mathrm{P}(X=0) \to 0$.

Proof. This is immediate by Chebyshev's inequality (2), applied to $X/g(Y)$.

It is also easy to see that the assumptions imply the assumptions of Theorem 1. In fact, let $Z = X/g(Y)$. Then $Z \to 1$ in L^2, and thus, cf. (17), $\mathrm{E}(\mathrm{Var}(Z \mid Y)) \to 0$ and $\mathrm{E}(Z \mid Y) \to 1$ in L^2 (and thus in probability). Hence $\mathrm{Var}(X \mid Y)/\mathrm{E}(X \mid Y)^2 = \mathrm{Var}(Z \mid Y)/\mathrm{E}(Z \mid Y)^2 \xrightarrow{p} 0$, and $\mathrm{P}(\mathrm{E}(X \mid Y) = 0) = \mathrm{P}(\mathrm{E}(Z \mid Y) = 0) \to 0$. □

Example 4. Janson [4] studied for example the number of spanning subtrees in $G(n,p)$, where p may depend on n. Denote this number by X and let Y be the number of edges. Then $\mathrm{E}\,X = n^{n-2}p^{n-1}$, which suggest an approximation of the form

$$\begin{aligned} c_1\Bigl(\frac{Y}{\binom{n}{2}}\Bigr)^{n-1} &= c_2\Bigl(\frac{Y}{\mathrm{E}\,Y}\Bigr)^{n-1} = c_2 e^{(n-1)\log(Y/\mathrm{E}\,Y)} \\ &\approx c_2 e^{(n-1)(Y/EY-1)} = c_3 e^{aY}, \end{aligned}$$

with $a = 2p/n$. It is convenient to change this a little, and we choose

$$g(Y) = c^{-1}(1-a)^{-Y}$$

with $c^{-1} = \mathrm{E}\bigl(X(1-a)^Y\bigr)$. Then $X/g(Y) = cX(1-a)^Y$ and $\mathrm{E}(X/g(Y)) = 1$. It turns out, cf. [4, (4.15) and Remark 4.1], that if $\limsup p < 1$ and $p \gg n^{-1/2}$, then $\mathrm{Var}(X/g(Y)) \to 0$. Hence Theorem 5 applies; moreover, $\log X - \log g(Y) \xrightarrow{p} 0$, which implies that X has the same asymptotic log-normal distribution as $g(Y)$.

Consider now the case $n^{1/2}p \to c > 0$. Then $\mathrm{Var}\,X/g(Y) \to \alpha$ for some $\alpha > 0$. However, it follows by the argument in [4, section 4] that if Y' denotes the number of paths of length 2 in the graph, then $\mathrm{Var}\bigl(X/g(Y) - h(Y')\bigr) \to 0$ for a suitable function h. Hence

$$\mathrm{Var}\bigl(X/g(Y)\bigr) = \mathrm{Var}\bigl(\mathrm{E}(X/g(Y) \mid Y')\bigr) + o(1),$$

and Theorem 2 applies to $X/g(Y)$ and Y', which again gives an asymptotic log-normal distribution of X. It follows that, in this case, X can be approximated (with small relative error) by a function of (Y, Y'). We conjecture that this is the case also when $n^{1/2}p \to 0$ as long as $n^{2/3}p \to \infty$; for smaller p we conjecture that X can be well approximated by a function of

(Y, Y', Y''), where Y'' is the number of subtrees of order 4 (paths of length 3 and 3-stars), as long as $n^{3/4}p \to \infty$, and so on with more and more terms for smaller p.

The discussion above applies also to some similar variables treated in [4], for example the number of Hamilton cycles and the number of perfect matchings. Note that the result in the latter case corresponds to the hypergraph result in Example 2, and that the condition $m \gg n^{4/3}$ corresponds to $p \gg n^{-2/3}$, which we conjectured to be the right condition for approximation by a function of Y and Y'. Note also that both Y and Y' are functions of the sequence of vertex degrees, so the conditioning used in Example 2 is stronger than conditioning on (Y, Y').

References

1. C. Cooper, A. Frieze, M. Molloy and B. Reed, Perfect matchings in random r-regular, s-uniform hypergraphs, *Combinatorics, Probability and Computing* (to appear).
2. A. Frieze and S. Janson, Perfect matchings in random s-uniform hypergraphs, preprint, Uppsala, 1994.
3. S. Janson, Random trees in a graph and trees in a random graph, *Math. Proc. Camb. Phil. Soc.* **100** (1986), 319–330.
4. S. Janson, The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph, *Combinatorics, Probability and Computing* **3** (1994) (to appear).
5. S. Janson, Random regular graphs: asymptotic distributions and contiguity, preprint, Minneapolis, 1993.
6. J. W. Moon, The second moment of the complexity of a graph, *Mathematika* **11** (1964), 95–98.
7. R.W. Robinson and N.C. Wormald, Almost all cubic graphs are hamiltonian, *Random Struct. Alg.* **3** (1992), 117–125.
8. R.W. Robinson and N.C. Wormald, Almost all regular graphs are hamiltonian, *Random Struct. Alg.* **5** (1994), 363–374.
9. J. Schmidt and E. Shamir, A treshold for perfect matchings in random d-pure hypergraphs, *Discrete Math.* **45** (1983), 287–295.

Svante Janson, Department of Mathematics, Uppsala University, PO Box 480, S-751 06 Uppsala, Sweden

E-mail address: `svante.janson@math.uu.se`