A Symbolic Computation Approach to a Problem Involving Multivariate Poisson Distributions (Draft, do not distribute!)*

Eduardo D. Sontag and Doron Zeilberger
Department of Mathematics
Rutgers University
Hill Center-Busch Campus
110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA.
[sontag,zeilberg] at math dot rutgers dot edu

May 5, 2009

1 Introduction

Eduardo, please write it.

******* important to decide: do we really start by emphasizing chemistry, or do we first state the purely mathematical problem (including Prussian officers and email messages. well, no Prussian officers, but we can make a good story with text messages, emails, and so on)? Then the chemistry would be a self-contained second part. DZ's reply: Yes, good idea! Keep the Chemistry to the second part, but mention it in the still-to-be-written abstart. *******

2 The Problem

Suppose that we have n independent Poisson random variables, X_j $(j = 1 \ldots n)$, with parameters λ_j respectively. In other words

$$\Pr\left(X_1 = k_1, X_2 = k_2, \ldots, X_n = k_n\right) = e^{-(\lambda_1 + \cdots + \lambda_n)} \frac{\lambda_1^{k_1}}{k_1!} \frac{\lambda_2^{k_2}}{k_2!} \cdots \frac{\lambda_n^{k_n}}{k_n!} \quad . \tag{1}$$

*Accompanied by Maple package MVPoisson downloadable from
http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/mvp.html

Suppose that we can't observe the X_j's directly, but only a certain number, m, of linear combinations of them:

$$Y_i = \sum_{j=1}^{n} a_{ij} X_j \quad , \quad (i = 1, \ldots, m) \quad ,$$

where (a_{ij}) is a certain $m \times n$ matrix with non-negative coefficients.

We are interested in the following questions:

1. Can one compute (fast!), for any given vector $(b_1, \ldots, b_m)$ (possibly with large coordinates), the probability

$$F(b_1, \ldots, b_m) := \Pr\left(Y_1 = b_1, \ldots, Y_m = b_m\right) \quad .$$

2. Can one compute (fast!), for any given vector $(b_1, \ldots, b_m)$, (possibly with large coordinates) the conditional expectation

$$G_j(b_1, \ldots, b_m) := E[X_j \mid Y_1 = b_1, \ldots, Y_m = b_m] \quad , \quad (1 \leq j \leq n).$$

3. More generally, can one compute (fast!), the higher moments

$$G_j^{(r)}(b_1, \ldots, b_m) := E[X_j^r \mid Y_1 = b_1, \ldots, Y_m = b_m] \quad , \quad (r \geq 2) \quad ,$$

that would immediately allow us to compute the moments about the mean. Can we compute (fast!) mixed moments, in particular the covariances?

******* EDS will fill-in: As trivial example, consider the case in which there is one column... multinomial *******

******* actually, we are also VERY interested in covariances, so we need JOINT moments too. Doron: could you write that up easily, or should we say something like "one can also easily find them" for now, if we want to be done soon? *******

3 The generating function

Fix a matrix $A = (a_{ij})$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$, once and for all. Let

$$F_0(b_1, \ldots, b_m) = \sum_{\substack{k_1, \ldots, k_n \geq 0 \\ a_{11}k_1 + \cdots + a_{1n}k_n = b_1, \ldots, a_{m1}k_1 + \cdots + a_{mn}k_n = b_m}} \frac{\lambda_1^{k_1}}{k_1!} \frac{\lambda_2^{k_2}}{k_2!} \cdots \frac{\lambda_n^{k_n}}{k_n!}$$

(value is zero if the sum is empty), so that

$$F(b_1, \ldots, b_m) = e^{-(\lambda_1 + \cdots + \lambda_n)} F_0(b_1, \ldots, b_m) \, .$$

Let f_0 be the (multivariable) generating function of F_0, in other words

$$f_0(z_1,\ldots,z_m)=\sum_{b_1\geq 0,\ldots,b_m\geq 0} F_0(b_1,\ldots,b_m) z_1^{b_1}\cdots z_m^{b_m} \quad .$$

Our quantity of interest, $F_0(b_1,\ldots,b_m)$, is the coefficient of $z_1^{b_1}\cdots z_m^{b_m}$ in the multivariable Taylor expansion about the origin of $f_0(z_1,\ldots,z_m)$.

We have:

$$f_0(z_1,\ldots,z_m)=\sum_{b_1\geq 0,\ldots,b_m\geq 0}\left(\sum_{\substack{k_1,\ldots,k_n\geq 0\\ a_{11}k_1+\cdots+a_{1n}k_n=b_1,\ldots,a_{m1}k_1+\cdots+a_{mn}k_n=b_m}} \frac{\lambda_1^{k_1}}{k_1!}\frac{\lambda_2^{k_2}}{k_2!}\cdots\frac{\lambda_n^{k_n}}{k_n!}\right) z_1^{b_1}\cdots z_m^{b_m} \quad .$$

By changing the order of summation, this equals

$$\begin{aligned}
&\sum_{k_1\geq 0,\ldots,k_n\geq 0} \frac{\lambda_1^{k_1}}{k_1!}\frac{\lambda_2^{k_2}}{k_2!}\cdots\frac{\lambda_n^{k_n}}{k_n!} z_1^{a_{11}k_1+\cdots+a_{1n}k_n}\cdots z_m^{a_{m1}k_1+\cdots+a_{mn}k_n}\\
&= \sum_{k_1\geq 0,\ldots,k_n\geq 0} \frac{(\lambda_1 z_1^{a_{11}} z_2^{a_{21}}\cdots z_m^{a_{m1}})^{k_1}}{k_1!}\cdots\frac{(\lambda_n z_1^{a_{1n}} z_2^{a_{2n}}\cdots z_m^{a_{mn}})^{k_n}}{k_n!}\\
&= \left(\sum_{k_1\geq 0}\frac{(\lambda_1 z_1^{a_{11}} z_2^{a_{21}}\cdots z_m^{a_{m1}})^{k_1}}{k_1!}\right)\cdots\left(\sum_{k_n\geq 0}\frac{(\lambda_n z_1^{a_{1n}} z_2^{a_{2n}}\cdots z_m^{a_{mn}})^{k_n}}{k_n!}\right)\\
&= \exp(\lambda_1 z_1^{a_{11}} z_2^{a_{21}}\cdots z_m^{a_{m1}})\cdots\exp(\lambda_n z_1^{a_{1n}} z_2^{a_{2n}}\cdots z_m^{a_{mn}})\\
&= \exp\left(\lambda_1 z_1^{a_{11}} z_2^{a_{21}}\cdots z_m^{a_{m1}}+\cdots+\lambda_n z_1^{a_{1n}} z_2^{a_{2n}}\cdots z_m^{a_{mn}}\right) \quad .
\end{aligned}$$

We have just derived

Theorem 1:

$$f_0(z_1,\ldots,z_m)=\exp\left(\sum_{j=1}^{n}\lambda_j\prod_{i=1}^{m} z_i^{a_{ij}}\right)$$

The conditional probability

$$\Pr\left(X_1=k_1, X_2=k_2,\ldots,X_n=k_n \;\middle|\; Y_1=b_1,\ldots,Y_m=b_m\right)$$

is the same as the expression in (1) divided by $F(b)$, provided that $\sum_{j=1}^{n} a_{ij}k_j=b_i$ for all i, and is zero otherwise. Recall that the rth factorial moment of a random variable W, $E[W^{(r)}]$, is, by definition, the expectation of $W!/(W-r)!$. We are interested in the conditional factorial moments of X_j given $Y=b$, which we will denote as $E[X_j^{(r)} \mid Y]$. By definition, $E[X_j^{(r)} \mid Y]$ the following expression and divided by $F_0(b)$:

$$\sum_{\substack{k_1,\ldots,k_n\geq 0\\ a_{11}k_1+\cdots+a_{1n}k_n=b_1,\ldots,a_{m1}k_1+\cdots+a_{mn}k_n=b_m}} k_j(k_j-1)\ldots(k_j-r+1)\frac{\lambda_1^{k_1}}{k_1!}\frac{\lambda_2^{k_2}}{k_2!}\cdots\frac{\lambda_n^{k_n}}{k_n!} \quad . \qquad (2)$$

Now, expression (2) is the same as the result of applying the operator $\lambda_j^r(\frac{\partial}{\partial\lambda_j})^r$ to $F_0(b_1,\ldots,b_m)$ *when viewing the λ's as variables and not as constants.* On the other hand,

$$\lambda_j^r(\frac{\partial}{\partial\lambda_j})^r f_0(z_1,\ldots,z_m)=\sum_{b_1\geq 0,\ldots,b_m\geq 0}\lambda_j^r(\frac{\partial}{\partial\lambda_j})^r F_0(b_1,\ldots,b_m) z_1^{b_1}\cdots z_m^{b_m}$$

and therefore expression (2) is the same as the coefficient of $z_1^{b_1} \cdots z_m^{b_m}$ in $\lambda_j^r (\frac{\partial}{\partial \lambda_j})^r f_0(z_1, \ldots, z_m)$ Since, as formal power series, we have the representation in Theorem 1. we conclude that expression (2) is the same as the coefficient of $z_1^{b_1} \cdots z_m^{b_m}$ in $(\prod_{i=1}^m z_i^{a_{ij}})^r f_0(z)$, which is the same as $F(b_1 - ra_{1j}, b_2 - ra_{2j}, \ldots, b_m - ra_{mj})$ when all $b_i - ra_{ij} \geq 0$ and zero otherwise. In conclusion, $E[X_j^{(r)} \mid Y]$ equals $F_0(b_1 - ra_{1j}, b_2 - ra_{2j}, \ldots, b_m - ra_{mj})$ divided by $F_0(b)$. We have proved:

Theorem 2: The conditional factorial moments $E[X_j^{(r)} \mid Y]$ are given in terms of the $F_0(b_1, \ldots, b_m)$ by

$$\lambda_j^r \cdot \frac{F_0(b_1 - ra_{1j}, b_2 - ra_{2j}, \ldots, b_m - ra_{mj})}{F_0(b_1, \ldots, b_m)}$$

when all $b_i - ra_{ij} \geq 0$ and zero otherwise.

So everything depends on a fast computation of the coefficients $F_0(b_1, \ldots, b_m)$, of $f_0(z_1, \ldots, z_m)$.

By taking mixed partial derivatives, we can easily derive analogous expressions for mixed moments, in particular, the covariances.

4 Recurrences

From now on, let's assume that the entries of A, (a_{ij}), are non-negative **integers**. In that case, we can write

******* next line is not quite right unless the exponentials are taken into account in the definition of Q. Or, equivalently, we can compute f_0. DZ: Indeed we should stick to f_0. *******

$$f_0(z) = \exp(Q(z_1, \ldots, z_m)) \quad ,$$

where $Q(z_1, \ldots, z_m)$ is the **polynomial**

$$Q(z_1, \ldots, z_m) := \sum_{j=1}^{n} \lambda_j \prod_{i=1}^{m} z_i^{a_{ij}} \quad .$$

By Cauchy's theorem, we can express $F(b_1, \ldots, b_m)$ as a **multi-contour integral**:

$$F(b_1, \ldots, b_m) = \left(\frac{1}{2\pi i}\right)^m \int_{|z_1|=c} \cdots \int_{|z_m|=c} \frac{\exp(Q(z_1, \ldots, z_m))}{z_1^{b_1+1} \cdots z_m^{b_m+1}} \, dz_1 \, \ldots \, dz_m \quad .$$

By the celebrated **Wilf-Zeilberger** theory ([WZ]), $F(b_1, \ldots, b_m)$ satisfies pure **linear recurrences with polynomial coefficients** in each of its arguments. This means that for each i between 1 and m, there exists a positive integer R_i (the order) and polynomials $P_r^{(i)}(b_1, \ldots, b_m)$ $(0 \leq r \leq R_i)$ such that the following holds, for *all* $(b_1, \ldots, b_m)$:

$$\sum_{r=0}^{R_i} P_r^{(i)}(b_1, \ldots, b_m) F(b_1, \ldots, b_{i-1}, b_i + r, b_{i+1}, \ldots, b_m) = 0 \quad .$$

Once these recurrences are known, one can compute $F(b_1, \ldots, b_m)$ in time linear in $b_1 + \cdots + b_m$ and with constant memory allocation (one only needs to remember, at each stage, a constant number of values).

******* add a warning about zeros of coefs of recursion? DZ: OK, here goes *******

In rare cases, the leading term of the recurrence would vanish, in which case, we would encounter a (discrete) "singularity", and would not be able to go on, since we would have to divide by 0, but in that case one can show that there is an alternative route, using another order of applying the recurrences.

The **Apagodu-Zeilberger**[ApZ] multi-variable extension of the Almkvist-Zeilberger[AlZ] algorithm can find such recurrences explicitly. Unfortunately, for matrices A with more than three rows, the time taken to find such recurrences is prohibitive, but many matrices of interest have two or three rows.

5 Two-Rowed matrices

If the matrix A only has two rows, and the entries are only $\{0, 1\}$, then one can express $F_(b_1, b_2)$ as a *single sum*. Indeed, let

- c_{01} be the sum of the λ_j's for which $a_{1,j} = 0$, $a_{2,j} = 1$,
- c_{01} be the sum of the λ_j's for which $a_{1,j} = 1$, $a_{2,j} = 0$,
- c_{01} be the sum of the λ_j's for which $a_{1,j} = 1$, $a_{2,j} = 1$.

Then, we have

$$Q(z) = c_{01}z_1 + c_{10}z_2 + c_{11}z_1z_2,$$

and so

$$\begin{aligned} f_0(z_1, z_2) &= e^{Q(z)} = \sum_{k=0}^{\infty} \frac{Q(z)^k}{k!} \\ &= \sum_{\alpha\geq 0,\beta\geq 0,\gamma\geq 0} \frac{(c_{01}z_1)^{\alpha}(c_{10}z_2)^{\beta}(c_{11}z_1z_2)^{\gamma}}{\alpha!\beta!\gamma!} \\ &= \sum_{\alpha\geq 0,\beta\geq 0,\gamma\geq 0} \frac{c_{01}^{\alpha}c_{10}^{\beta}c_{11}^{\gamma}z_1^{\alpha+\gamma}z_2^{\beta+\gamma}}{\alpha!\beta!\gamma!} \quad . \end{aligned}$$

To get $F_0(b_1, b_2)$, we must extract the coefficient of $z_1^{b_1}z_2^{b_2}$ which entails $\alpha = b_1 - \gamma, \beta = b_2 - \gamma$, and we have the single-sum binomial coefficient (hypergeometric) sum (replacing γ by k)

$$F_0(b_1, b_2) = \sum_{k=0}^{\min(b_1,b_2)} \frac{c_{11}^{k}c_{01}^{b_1-k}c_{10}^{b_2-k}}{k!(b_1-k)!(b_2-k)!} \quad .$$

Using the **Zeilberger Algorithm** ([Z, PWZ]) , we get the following linear recurrence:

$$(c_{10}b_1^2 + 4c_{10} - 2c_{10}b_2 + 4c_{10}b_1 - c_{10}b_1b_2)F_0(b_1 + 2, b_2)+$$

$$(-c_{11}b_1 - c_{11} + b_2c_{11} + b_2c_{10}c_{01} - 2b_1c_{10}c_{01} - 3c_{01}c_{10})F_0(b_1+1, b_2) + (c_{11}c_{10} + c_{01}c_{10}^2)F_0(b_1, b_2) = 0 \quad .$$

6 The Maple package MVPoisson

All this is implemented in the Maple package `MVPoisson` accompanying this article. It is available from the webpage of this article

`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/mvp.html` ,

where one can also find sample input and output.

******* paper: I looked at the sample outputs on the web, and saw statements such as:**

"This seems to be asymptotically etc etc..."

Can you please include a sentence explaining how this fit is done by the software? Or did you do this by hand? Is there any reason to expect $\Omega(n) = cn^p$? *****

******* I'd like to download all at once all the files linked from**

http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/mvp.html, but for some strange reason my "wget" is failing to work recursively (you must have a funny directory structure). Is it possible to get a tar file? *****

References

[AlZ] G. Almkvist and D. Zeilberger, *The method of differentiating under the integral sign*, J. Symbolic Computation **10**(1990), 571-591.

[ApZ] M. Apagodu and D. Zeilberger, *Multi-Variable Zeilberger and Almkvist-Zeilberger Algorithms and the Sharpening of Wilf-Zeilberger Theory* , Adv. Appl. Math. **37** (2006)(Special Regev issue), 139-152

[PWZ] M. Petkovsek, H.S. Wilf and D. Zeilberger, *A=B*, AK Peters, Wellesley, (1996). [available on-line from the authors' websites.]

[WZ] H.S. Wilf and D. Zeilberger, *An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities*, Invent. Math. **108** (1992), 575-633.

[Z] D. Zeilberger, *The method of creative telescoping*, J. Symbolic Computat. **11**, 195-204 (1991).