Conditional moments, gamma, free gamma, and free Poisson laws

Włodzimierz Bryc

Mahdia, June 1, 2005

Abstract This talk is based on joint paper with M. Bożejko "On a class of free Lévy laws related to a regression problem".

1 Univariate gamma law

Suppose X, Y > 0 are **non-degenerate** independent random variables. Let S = X + Y be their sum and $Z = \frac{X}{S}$ be the quotient.

Theorem 1 ([Lukacs, 1955]) If S and Z are independent, then X is gamma with density $\frac{1}{\Gamma(p)}x^{p-1}e^{-x}$, x > 0, p > 0 after normalization.

Simple proof Use conditional moments. If Z and S are

Slide 3

independent then E(X|S) = E(ZS|S) = SE(Z) and $E(X^2|S) = E(Z^2S^2|S) = S^2E(Z^2)$

Thus $E(X|S) = \mu S$ and $E(X^2|S) = (\sigma^2 + \mu^2)S^2$. Or

 $E(X|S) = \mu S$ and $Var(X|S) = \sigma^2 S^2$



Can we determine the distribution of X from E(X|S) and Var(X|S)?

Yes, as noticed in [Wesołowski, 1989].

Theorem 2 [Laha and Lukacs, 1960] Suppose X, Y are independent, E(X) = E(Y) = 0, $E(X^2) = E(Y^2) = 1$, S = X + Y.

$$E(X|S) = \frac{1}{2}S$$

and for some constants C, a, b

 $\operatorname{Var}(X|S) = C(1 + \frac{a}{2}S + \frac{b}{4}S^2).$

Then X and Y have the classical Meixner type law. In particular, C = 1/(2+b) and X is ...



Simplest proof? For any i.i.d. matrices, $E(\mathbf{X}|\mathbf{S}) = \frac{1}{2}\mathbf{S}$ If \mathbf{X}, \mathbf{Y} are independent Wishart matrices with shape parameters p, q, then [Letac and Massam, 1998] show that there are a = a(p,q), b = b(p,q) such that $\operatorname{Var}(\mathbf{X}|\mathbf{S}) = a(\operatorname{tr}\mathbf{S})\mathbf{S} + b\mathbf{S}^2$ The "simplest proof" fails. A "simple proof" in [Letac and Massam, 1998] relies on the quadratic regression property of **other** quadratic functions of \mathbf{X} .

2 Matrix gamma: Wishart law
Suppose \mathbf{X}, \mathbf{Y} are non-degenerate independent symmetric semi-positive-definite matrices. Consider
$\mathbf{S} = \mathbf{X} + \mathbf{Y}, \ \mathbf{Z} = \mathbf{S}^{-1/2} \mathbf{X} \mathbf{S}^{-1/2}$
[Olkin and Rubin, 1962], [Casalis and Letac, 1996], [Bobecka and Wesołowski, 2002] prove
Theorem 3 If $\mathbf{S} > 0$, \mathbf{X}, \mathbf{Y} are not concentrated on the same one-dimensional subspace, the law of \mathbf{Z} is invariant under orthogonal transformations, or \mathbf{X}, \mathbf{Y} have strictly positive twice-differentiable densities, and \mathbf{Z} and \mathbf{S} are independent, then \mathbf{X} is Wishart, $E(\exp\langle\theta, \mathbf{X}\rangle) = \det(\mathbf{I} - \theta)^{-p}$, $p > \frac{n-1}{2}$, after re-normalization.

ide 6

2.1 Large Wishart matrices
$\mathbf{X}_n, \mathbf{Y}_n$ are i.i.d. $n \times n$ Wishart matrices with shape parameter
$p > (n-1)/2$. $\mathbf{S}_n = \mathbf{X}_n + \mathbf{Y}_n$.
Goal: What can we say about the limit as $n \to \infty$, $p \to \infty$,
$p/n \to \lambda/2 > 0?$
$E(\frac{\mathbf{X}_n}{n} \frac{\mathbf{S}_n}{n}) = \frac{1}{2}\frac{\mathbf{S}_n}{n}$ in the limit gives $\mathcal{E}(\mathbb{X} \mathbb{S}) = \frac{1}{2}\mathbb{S}$
$E\left(\left(\frac{\mathbf{X}_n}{n}\right)^2 \middle \frac{\mathbf{S}_n}{n}\right) = \frac{np}{16p^2 + 4p - 2} \frac{\mathbf{S}_n}{n} \operatorname{tr}_n \frac{\mathbf{S}_n}{n} + \frac{4p^2 + 2p - 1}{16p^2 + 4p - 2} \left(\frac{\mathbf{S}_n}{n}\right)^2$
$E(\exp(\theta, \mathbf{S}_n)) = \det(I - \theta)^{-2p}$, so $E \exp \alpha \operatorname{tr}_n \frac{\mathbf{S}_n}{\mathbf{r}} = (1 - \alpha/n^2)^{-2pn}$.
So $\operatorname{tr}_n\left(\frac{\mathbf{S}_n}{n}\right) \to \lambda$ in prob.
$\operatorname{Var}(\mathbb{X} \mathbb{S}) = \frac{1}{8}\mathbb{S}, \text{ or } \operatorname{Var}(2\mathbb{X} 2\mathbb{S}) = \frac{1}{4}(2\mathbb{S})$
Slide 5: Conditional variance of 2X is like Poisson, not like gamma!

Slide 8

3 What are non-commutative r.v.?

Self-adjoint elements $\mathbb{X} = \mathbb{X}^*$ of a complex *-algebra \mathcal{A} with identity; preferably, von Neumann algebra.

State £ : A → C. Faithful. Tracial; preferably, normal normalized positive linear functional: E(a^{*}) = E(a), E(I) = 1, E(aa^{*}) ≥ 0. E(aa^{*}) = 0 implies a = 0. E(ab) = E(ba); preferably, continuous in weak*-topology.

Slide 11

• Law of X: probability measure μ such that $\mathcal{E}(\mathbb{X}^n) = \int_{\mathbb{R}} x^n \mu(dx)$

Example \mathcal{A} =random $n \times n$ matrices with $\mathcal{E}(a) = E(\operatorname{tr}_n(a))$. $\mathbb{X}, \mathbb{Y}, \ldots$ are random Hermitian matrices Voiculescu's theorem [Dykema, 1993] says that as $n \to \infty$ Hermitian matrices $\mathbb{X}/n, \mathbb{Y}/n$ with i.i.d entries are asymptotically free. [Capitaine and Casalis, 2004] show that independent Wishart

matrices are asymptotically free. Note: XY is not a r.v.!

3.2 Conditional expectations

Let $\mathcal{B} \subset \mathcal{A}$ be a *-subalgebra. The conditional expectation is a linear map $\mathcal{E}_{\mathcal{B}} : \mathcal{A} \to \mathcal{B}$ such that $\mathcal{E}_{\mathcal{B}}(\mathbb{Y}_1 \mathbb{X} \mathbb{Y}_2) = \mathbb{Y}_1 \mathcal{E}_{\mathcal{B}}(\mathbb{X}) \mathbb{Y}_2$ for all $\mathbb{X} \in \mathcal{A}, \mathbb{Y}_1, \mathbb{Y}_2 \in \mathcal{B}$. Note: here \mathbb{X}, \mathbb{Y} are not r.v.! Properties: Probabilistic notation: $\mathcal{E}_{\mathcal{B}}(\mathbb{X}) = \mathcal{E}(\mathbb{X}|\mathcal{B})$.

- (i) If $\mathbb{X} \in \mathcal{A}, \mathbb{Y} \in \mathcal{B} \subset \mathcal{A}$, then $\mathcal{E}(\mathbb{X}\mathbb{Y}) = \mathcal{E}(\mathcal{E}(\mathbb{X}|\mathcal{B})\mathbb{Y})$
- (ii) If random variables $\mathbb{U}, \mathbb{V} \in \mathcal{A}$ are free then $\mathcal{E}(\mathbb{U}|\mathbb{V}) = \mathcal{E}(\mathbb{U})\mathbb{I}$.
- (iii) Let \mathbb{W} be a (self-adjoint) element of the von Neumann algebra $\mathcal{B}_{\mathbb{V}}$ generated by a self-adjoint $\mathbb{V} \in \mathcal{A}$. If for all $n \geq 1$ we have $\mathcal{E}(\mathbb{U}\mathbb{V}^n) = \mathcal{E}(\mathbb{W}\mathbb{V}^n)$ then $\mathcal{E}(\mathbb{U}|\mathbb{V}) = \mathbb{W}$.

(iv) If $\mathcal{E}(\mathbb{U}_1\mathbb{V}^n) = \mathcal{E}(\mathbb{U}_2\mathbb{V}^n)$ for all $n \ge 1$, then $\mathcal{E}(\mathbb{U}_1|\mathbb{V}) = \mathcal{E}(\mathbb{U}_2|\mathbb{V})$.



Proof

(i) See [Takesaki, 1972]. (iv) Apply (iii). (ii) If Z is in the von Neumann algebra generated by V, then
\$\mathcal{E}((U - cI)Z) = \mathcal{E}(U - cI)\mathcal{E}(Z)\$. Applying this to
\$\mathbb{Z} = \mathcal{E}(U|V) - \mathcal{E}(U)I and \$c = \mathcal{E}(U)\$ after taking into account (i) we get \$\mathcal{E}(Z^2) = \mathcal{E}(Z(\mathcal{E}(U|V) - cI)) = \mathcal{E}(Z(U - cI|V)) = \$\mathcal{E}(Z(U - cI)) = \mathcal{E}(Z(U - cI)) = \mathcal{E}(Z(U - cI)) = \mathcal{E}(Z(U - cI)) = 0\$. Thus \$\mathcal{E}(U|V) = \mathcal{E}(U)I\$.
(iii) Let \$\mathbb{W}' = \mathcal{E}(U|V)\$. Then \$\mathcal{E}((W - \mathbb{W}')p(V)) = 0\$ for all polynomials \$p\$. Since polynomials \$p(V)\$ are dense in the von Neumann algebra generated by V, and \$\mathcal{E}(\cdot)\$ is normal, this implies that \$\mathcal{E}((W - \mathbb{W}')(\mathcal{W} - \mathbb{W}')^*) = 0\$; by faithfulness of \$\mathcal{E}(\cdot)\$ we deduce that \$\mathbb{W}' = \mathcal{W}\$.

4 Free version of Olkin-Rubin Theorem
Proposition 4 ([Bożejko and Bryc, 2004]) Suppose random variables X, Y ∈ A are free, identically distributed, σ = √Var(X) > 0, and such that S = X + Y is strictly positive; in particular, m = E(X) > 0. Let Z = S^{-1/2}XS^{-1/2}. If Z and S are free, then X has free-Poisson type law µ_{a,0} with a = σ/m.
For converse, see [Capitaine and Casalis, 2004, Corollary 7.2].
Simple proof
By exchangeability, E(X|S) = S/2. Also E(Z) = 1/2, as E(Z) = E(S^{-1/2}YS^{-1/2}) and E(S^{-1/2}(X + Y)S^{-1/2}) = E(I) = 1.
We now verify that Var(X|S) is a linear function of S. Denote the centering operation by U° = U – E(U)I.

le 13

le 14

Therefore,

Slide 15

Slide 16

$$\mathcal{E}(\mathbb{X}^2 \mathbb{S}^m) = \mathcal{E}\left(\left(\frac{1}{4}\mathbb{S}^2 + 2m \operatorname{Var}(\mathbb{Z})\mathbb{S}\right)\mathbb{S}^m\right),\,$$

which by Slide 11 (iii) implies that

$$\mathcal{E}(\mathbb{X}^2|\mathbb{S}) = \frac{1}{4}\mathbb{S}^2 + 2m\operatorname{Var}(\mathbb{Z})(\mathbb{S} - 2m) + 4m^2\operatorname{Var}(\mathbb{Z})\mathbb{I}.$$

Passing to standardized random variables $\mathbb{X}^{\circ}/\sigma$, $\mathbb{Y}^{\circ}/\sigma$, we get

$$\operatorname{Var}(\frac{1}{\sigma}\mathbb{X}^{\circ}|\mathbb{S}) = \frac{m^{2}\operatorname{Var}(\mathbb{Z})}{\sigma^{2}}\left(4\mathbb{I} + 2\frac{\sigma}{m}\mathbb{S}^{\circ}/\sigma\right).$$

This shows that standardized random variables satisfy

$$\operatorname{Var}(\mathbb{X}_s|\mathbb{S}) = C(\mathbb{I} + \frac{a}{2}\mathbb{S}_s)$$

with $a = \sigma/m$. (This also determines $\operatorname{Var}(\mathbb{Z}) = \sigma^2/(8m^2)$.)

Using tracial property and freeness of \mathbb{S}, \mathbb{Z} : $\underline{\mathcal{E}(\mathbb{X}^2 \mathbb{S}^m)} = \mathcal{E}(\mathbb{Z} \mathbb{S} \mathbb{Z} \mathbb{S}^{m+1}) = \mathcal{E}(\mathbb{Z} \mathbb{S} (\mathbb{Z}^\circ + 1/2\mathbb{I}) \mathbb{S}^{m+1})$ $= \frac{1}{2} \mathcal{E}(\mathbb{Z} \mathbb{S}^{m+2}) + \mathcal{E}(\mathbb{Z} (\mathbb{S}^\circ + 2m\mathbb{I}) \mathbb{Z}^\circ \mathbb{S}^{m+1})$ $= \frac{1}{4} \mathcal{E}(\mathbb{S}^{m+2}) + 2m \mathcal{E}(\mathbb{Z} \mathbb{Z}^\circ \mathbb{S}^{m+1}) + \mathcal{E}(\mathbb{Z} \mathbb{S}^\circ \mathbb{Z}^\circ \mathbb{S}^{m+1})$ $= \frac{1}{4} \mathcal{E}(\mathbb{S}^{m+2}) + 2m \operatorname{Var}(\mathbb{Z}) \mathcal{E}(\mathbb{S}^{m+1}) + \mathcal{E}(\mathbb{Z} \mathbb{S}^\circ \mathbb{Z}^\circ \mathbb{S}^{m+1}).$ The last term vanishes by freeness: $\mathcal{E}(\mathbb{Z} \mathbb{S}^\circ \mathbb{Z}^\circ \mathbb{S}^{m+1}) = \mathcal{E}((\mathbb{Z}^\circ + 1/2\mathbb{I}) \mathbb{S}^\circ \mathbb{Z}^\circ \mathbb{S}^{m+1})$ $= 1/2 \mathcal{E}(\mathbb{Z}^\circ) \mathcal{E}(\mathbb{S}^{m+2}) + \mathcal{E}(\mathbb{Z}^\circ \mathbb{S}^\circ \mathbb{Z}^\circ \mathbb{S}^{m+1})$ $= 0 + \mathcal{E}(\mathbb{Z}^\circ \mathbb{S}^\circ \mathbb{Z}^\circ) \mathcal{E}(\mathbb{S}^{m+1}) = 0.$

Can we determine the distribution of X from $\mathcal{E}(X|S)$ and Var(X|S)?

Theorem 5 ([Bożejko and Bryc, 2004]) Suppose $\mathbb{X}, \mathbb{Y} \in \mathcal{A}$ are free, self-adjoint, $\mathcal{E}(\mathbb{X}) = \mathcal{E}(\mathbb{Y}) = 0$, $\mathcal{E}(\mathbb{X}^2) = \mathcal{E}(\mathbb{Y}^2) = 1$, $\mathbb{S} = \mathbb{X} + \mathbb{Y}$ and

$$\mathcal{E}(\mathbb{X}|\mathbb{S}) = \frac{1}{2}\mathbb{S}$$
(1)

and there are numbers $C, a, b \in \mathbb{R}$ such that

$$\operatorname{Var}(\mathbb{X}|\mathbb{S}) = C(\mathbb{I} + \frac{a}{2}\mathbb{S} + \frac{b}{4}\mathbb{S}^2).$$
(2)

Then X and Y have the free Meixner law. In particular, $b \ge -1$, C = 1/(2+b) and the law of X is ...



5 Free Olkin-Rubin Theorem II

Slide 19

Proposition 6 ([Bożejko and Bryc, 2004]) Suppose $\mathbb{X}, \mathbb{Y} \in \mathcal{A}$ are free, identically distributed, $\sigma = \sqrt{\operatorname{Var}(\mathbb{X})} > 0$ and such that $\mathbb{S} = \mathbb{X} + \mathbb{Y}$ is strictly positive; in particular, $m = \mathcal{E}(\mathbb{X}) > 0$. Let $\mathbb{Z} = \mathbb{S}^{-1}\mathbb{X}^2\mathbb{S}^{-1}$. If \mathbb{Z} and \mathbb{S} are free, then \mathbb{X} has free-gamma type law μ_{2a,a^2} with $a = \sigma/m$.

Simple proof: By exchangeability, $\mathcal{E}(\mathbb{X}|\mathbb{S}) = \mathbb{S}/2$. By freeness, $\mathcal{E}(\mathbb{X}^2|\mathbb{S}) = \mathbb{S}\mathcal{E}(\mathbb{Z}|\mathbb{S})\mathbb{S} = \mathcal{E}(\mathbb{Z})\mathbb{S}^2$. Thus

$$\operatorname{Var}(\mathbb{X}|\mathbb{S}) = c\mathbb{S}^2,$$

where $c = \mathcal{E}(\mathbb{Z}) - 1/4 \ge 0$. After standardization Slide 16 (2) holds with $a = 2\sigma/m$, $b = \sigma^2/m^2$. (And $c = \sigma^2/(2m^2 + \sigma^2)$.)



6 Questions and Speculations

- (i) Why n = 1 and $n = \infty$ are simpler?
- (ii) If \mathbb{X}, \mathbb{Y} are free gamma, $\mathbb{S} = \mathbb{X} + \mathbb{Y}$, are \mathbb{S} and $\mathbb{S}^{-1}\mathbb{X}^2\mathbb{S}^{-1}$ indeed free?
- (iii) Does the matrix version of Olkin-Rubin II hold? Are there nontrivial i.i.d. $n \times n$ symmetric random matrices \mathbf{X}, \mathbf{Y} with independent sum $\mathbf{S} = \mathbf{X} + \mathbf{Y}$ and quotient $\mathbf{Z} = \mathbf{S}^{-1}\mathbf{X}^2\mathbf{S}^{-1}$?
- (iv) Is there a matrix version of [Laha and Lukacs, 1960]?
- (v) [Laha and Lukacs, 1960] holds in more generality. Does the free version generalize, too?

Question: Are there i.i.d. symmetric random matrices \mathbf{X}, \mathbf{Y} with independent $\mathbf{S} = \mathbf{X} + \mathbf{Y}$ and $\mathbf{Z} = \mathbf{S}^{-1}\mathbf{X}^2\mathbf{S}^{-1}$?
p%-trivial answers:
100% Let $\mathbf{X} = \text{diag}(X_1,, X_n)$ where X_j are independent gamma
 95% A distribution of X is a 95%-trivial answer to the query, if one can produce from it a new answer by taking X' = UXU* for a fixed deterministic orthogonal matrix U. Example: The Wishart distribution on the Lorenz cone i.e.
$\mathbf{X} = \begin{bmatrix} x & y \\ y & x \end{bmatrix}$ defined on $\Omega = \mathbb{R}^2$ with the density $f(x, y) = Ce^{-ax-by}(x^2 - y^2)^{p-1}$, where $x > y $, $p > 0$, $a, b > 0$ is a 95%-trivial answer.
0% Invariant under orthogonal transformations.

le 21

le 22

References [Bobecka and Wesołowski, 2002] Bobecka, K. and Wesołowski, J. (2002). The Lukacs-Olkin-Rubin theorem without invariance of the "quotient". Studia Math., 152(2):147-160. [Bożejko and Bryc, 2004] Bożejko, M. and Bryc, W. (2004). On a class of free Lévy laws related to a regression problem. arxiv.org/abs/math.OA/0410601. [Capitaine and Casalis, 2004] Capitaine, M. and Casalis, M. (2004). Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to beta random matrices. Indiana Univ. Math. J., 53(2):397-431. [Casalis and Letac, 1996] Casalis, M. and Letac, G. (1996). The Lukacs-Olkin-Rubin characterization of Wishart distributions on symmetric cones. Ann. Statist., 24(2):763-786. [Dykema, 1993] Dykema, K. (1993). On certain free product factors via an extended matrix model. J. Funct. Anal., 112(1):31-60. [Laha and Lukacs, 1960] Laha, R. G. and Lukacs, E. (1960). On a problem connected with quadratic regression. Biometrika, 47(300):335-343. [Letac and Massam, 1998] Letac, G. and Massam, H. (1998). Quadratic and inverse regressions for Wishart distributions. Ann. Statist., 26(2):573-595. [Lukacs, 1955] Lukacs, E. (1955). A characterization of the gamma distribution. Ann. Math. Statist., 26:319-324. [Morris, 1982] Morris, C. N. (1982). Natural exponential families with quadratic variance functions. Ann. Statist., 10(1):65-80. [Olkin and Rubin, 1962] Olkin, I. and Rubin, H. (1962). A characterization of the Wishart distribution Ann Math Statist 33:1272-1280 [Saitoh and Yoshida, 2001] Saitoh, N. and Yoshida, H. (2001). The infinite divisibility and orthogonal polynomials with a constant recursion formula in free probability theory. Probab. Math. Statist., 21(1):159-170.

Slide 23

Slide 24

Theorem 7 ([Laha and Lukacs, 1960]) Suppose X, Y are independent, E(X) = E(Y) = 0, $E(X^2) > 0$, $E(Y^2) > 0$, S = X + Y, and for some constants C, a, b, ρ $E(X|S) = \rho S$, $Var(X|S) = C(1 + \frac{a}{2}S + \frac{b}{4}S^2)$. Then X and Y have the classical Meixner type law. In particular, X is as in Slide 4. Question: Does the free version hold in more generality, too? If X, Y are free, non-degenerate, $\mathcal{E}(X|S) = \rho S$, $Var(X|S) = C(1 + aS + bS^2)$, does the analogous six-part conclusion from Slide 17 follow?

 [Takesaki, 1972] Takesaki, M. (1972). Conditional expectations in von Neumann algebras. J Functional Analysis, 9:306-321. [Voiculescu, 2000] Voiculescu, D. (2000). Lectures on free probability theory. In Lectures on probability theory and statistics (Saint-Flour, 1998), volume 1738 of Lecture Notes in Math., pages 279-349. Springer, Berlin. [Wesołowski, 1989] Wesołowski, J. (1989). Characterizations of some processes by properties conditional moments. Demonstratio Math., 22(2):537-556. 	J.
 [Voiculescu, 2000] Voiculescu, D. (2000). Lectures on free probability theory. In Lectures on probability theory and statistics (Saint-Flour, 1998), volume 1738 of Lecture Notes in Math., pages 279-349. Springer, Berlin. [Wesołowski, 1989] Wesołowski, J. (1989). Characterizations of some processes by properties conditional moments. Demonstratio Math., 22(2):537-556. 	
[Wesołowski, 1989] Wesołowski, J. (1989). Characterizations of some processes by properties conditional moments. Demonstratio Math., 22(2):537-556.	
	es of