mathoverflow

# A remarkable identity involving $\chi^2$ random variables

Asked 1 year, 3 months ago Modified 3 months ago Viewed 712 times

11

4

In the process of computing inclusion constants for the complex matrix cube (which is a free spectrahedron), the following identity was proven: for all $n \geq 1$,

$$\mathbb{E}\Big|\sum_{i=1}^{2n} x_i^2 - \sum_{j=1}^{2n} y_j^2\Big| = \mathbb{E}\Big|\sum_{i=1}^{2n} x_i^2 - \sum_{j=1}^{2n-2} y_j^2\Big| = 4^{1-n} n \binom{2n}{n},$$

where $x_i$,$y_j$ are i.i.d. standard real Gaussian random variables. The proof we have at this moment is by using the explicit form of the density of the difference of two $\chi^2$ random variables, see here and also *Klar, Bernhard*, **A note on gamma difference distributions**, ZBL07183251.

Since the result is so simple, there should be a more direct and more insightful proof of it.

**Question 1:** give an easy, conceptual proof of the identity above.

Consider now the function

$$k \mapsto \mathbb{E}\Big|\sum_{i=1}^{2k} x_i^2 - \sum_{j=1}^{2(n-k)} y_j^2\Big|$$

from $\{0, 1, \ldots, n\} \to \mathbb{R}_+$; as above, the $x_i$ and $y_j$ are standard i.i.d. Gaussians.

**Question 2:** Show that the function above is unimodular, and that its minimum is attained at $k = \lfloor n/2 \rfloor$.

There is a pretty involved proof of the second fact above in *Helton, J. William; Klep, Igor; McCullough, Scott; Schweighofer, Markus*, **Dilations, linear matrix inequalities, the matrix cube problem and beta distributions**, Memoirs of the American Mathematical Society 1232. Providence, RI: American Mathematical Society (AMS) (ISBN 978-1-4704-3455-7/pbk; 978-1-4704-4947-6/ebook). vi, 106 p. (2019). ZBL1447.47009.

**Question 3:** Does the claim in ``Question 2'' hold for arbitrary probability distributions?

Any help or insight about these questions would be appreciated!

co.combinatorics st.statistics probability-distributions

Share Cite Improve this question Follow

edited May 16 at 16:12 Jukka Kohonen

asked May 21, 2021 at 8:12 Ion Nechita

Here's the paper where the question popped out: arxiv.org/abs/2105.11302 see Lemma 6.2 and Theorem 6.3 – Ion Nechita May 25, 2021 at 8:16

2 Answers

Sorted by: Highest score (default)

10

I think I found an elementary proof of Question 2/3 for arbitrary probability distributions. In fact, it is not required that the components in the sums are squares, but general i.i.d. non-negative random variables work. Further, the requirement that both sums have an even number of terms ($2k$ and $2(n-k)$ in the question) is not required.

Let $X_1, \dots, X_N$ be i.i.d. non-negative, integrable random variables. Let $T_k := \mathbb{E}[|\sum_{i=1}^k X_i - \sum_{i=k+1}^N X_i|]$.

**Lemma 1:** For $k \geq \lfloor N/2 \rfloor$ it holds $T_{k+1} \geq T_k$.

By symmetry (for $k \leq \lfloor N/2 \rfloor$) this Lemma yields the question. To prove Lemma 1, we require the two following supplementary statements

**Lemma 2:** Let $C$ be a symmetric, integrable random variable, and $a, b \in \mathbb{R}$, $|b| \geq |a|$. Then $\mathbb{E}[|C+b|] \geq \mathbb{E}[|C+a|]$.

Hereby, $C$ being symmetric means that $C$ and $-C$ have the same distribution.

**Lemma 3:** Let $C$ be a symmetric, integrable random variable and $A$, $B$ non-negative and integrable random variables and assume that $A, B, C$ are independent. Then $\mathbb{E}[|C+A-B|] \leq \mathbb{E}[|C+A+B|]$.

***Proof of Lemma 2:*** Without loss of generality, by symmetry of $C$, let $0 \leq a \leq b$. Then one calculates

$$\begin{aligned}\mathbb{E}[|C+a|] &= \mathbb{E}[1_{a\geq|C|}(C+a) + 1_{a<|C|}(1_{C>0} + 1_{C<0})|C+a|] \\ &= \mathbb{P}(|C| \leq a)a + \mathbb{E}[1_{a<|C|}(1_{C>0}(C+a) + 1_{C<0}(-C-a)] \\ &= \mathbb{P}(|C| \leq a)a + \mathbb{E}[1_{|C|>a}|C|] \\ &= \mathbb{E}[\max\{a, |C|\}]\end{aligned}$$

and this term is obviously increasing in $a$. (in general, it simply holds $\mathbb{E}[|C+a|] = \mathbb{E}[\max\{|C|, |a|\}]$, which can also be proved via the identity $|C+a| + |-C+a| = 2\max\{|C|, |a|\}$ as pointed out by Fedor.)

***Proof of Lemma 3:*** Let $C \sim \mu, A \sim \nu, B \sim \theta$. Then

$$\mathbb{E}[|C+A-B| - |C+A+B|] = \int\int\Big(\underbrace{\int |c+a-b| - |c+a+b|\mu(dc)}_{\leq 0}\Big)\nu(da)\theta(db$$

since the term inbetween the large brackets is non-positive since $|a+b| \geq |a-b|$ (since $a, b \geq 0$) and by Lemma 2.

***Proof of Lemma 1:*** Let $C := \sum_{i=1}^{N-k-1} X_i - \sum_{i=k+1}^{N} X_i$ and note that $C$ is symmetric, $A := \sum_{i=N-k}^{k} X_i$ (where the sum is understood to be 0 if $k < N-k$), and $B := X_{k+1}$. Note that, for $k \geq \lfloor \frac{N}{2} \rfloor$, it holds $T_k = \mathbb{E}[|C + A - B|]$ and $T_{k+1} = \mathbb{E}[|C + A + B|]$. The claim now follows from Lemma 3.

Share Cite Improve this answer Follow

edited Jun 6, 2021 at 20:19

answered May 23, 2021 at 17:57

Steve

2 Nice. Lemma 2 follows just from an identity $|C+a| + |-C+a| = 2\max(|C|, |a|)$. – Fedor Petrov May 23, 2021 at 19:52

Beautiful proof, thanks a lot, that's exactly what I was looking for. Thanks Fedor for simplifying Lemma 2 – Ion Nechita May 24, 2021 at 8:11

---

16

1. The proof of

$$\mathbb{E}\Big|\sum_{i=1}^{2n} x_i^2 - \sum_{j=1}^{2n} y_j^2\Big| = 4^{1-n} n \binom{2n}{n}.$$

Denote $z_i = (x_i - y_i)/\sqrt{2}$, $w_i = (x_i + y_i)/\sqrt{2}$. Then the vectors $z$ and $w$ are i.i.d. standard Gaussian, and

$$\left|\sum_{i=1}^{2n} (x_i^2 - y_i^2)\right| = 2\,|\langle z, w\rangle|\,.$$

For finding the expectation of $|\langle z, w\rangle|$ we may fix $z$, if $|z| = a$, then $|\langle z, w\rangle|$ is distributed as $a|X|$ for standard Gaussian $X$, the expected value over $w$ is $a\sqrt{2/\pi}$ (I use that mean absolute value of $X$ equals $\sqrt{2/\pi}$). So, it remains to compute $\mathbb{E}|z|$ which is also well known.

2. The proof of

$$\mathbb{E}\Big|\sum_{k=1}^{2n} x_k^2 - \sum_{j=1}^{2n} y_j^2\Big| = \mathbb{E}\Big|\sum_{k=1}^{2n} x_k^2 - \sum_{j=1}^{2n-2} y_j^2\Big|.$$

We apply the integral representation

$$\frac{\pi}{2}(|b| - |a|) = \int_0^\infty \frac{\cos at - \cos bt}{t^2} dt.$$

Thus by Fubini theorem it suffices to prove that for each specific $t$ we have

$$\mathbb{E}\cos t\Big(\sum_{k=1}^{2n} x_k^2 - \sum_{j=1}^{2n} y_j^2\Big) = \mathbb{E}\cos t\Big(\sum_{k=1}^{2n} x_k^2 - \sum_{j=1}^{2n-2} y_j^2\Big). \qquad (\heartsuit)$$

For this we write $2\cos x = e^{ix} + e^{-ix}$ and apply the independence of $x_k$'s and $y_j$'s. We have for real $t$ and standard gaussian $X$

$$\mathbb{E}e^{itX^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itX^2 - X^2/2} dt = \frac{1}{\sqrt{1-2it}}$$

(the square root branch in the right half-plane is natural). Therefore $(\heartsuit)$ reads as

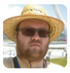
$$\frac{2}{(1-2it)^n(1+2it)^n} = \frac{1}{(1-2it)^n(1+2it)^{n-1}} + \frac{1}{(1-2it)^{n-1}(1+2it)^n}$$

that is true.

Share Cite Improve this answer Follow

edited May 22, 2021 at 7:19

answered May 21, 2021 at 11:41



Fedor Petrov

I would have loved to split the "accepted answer" token between yours and Steve's, but unfortunately MO enforces a choice :) – Ion Nechita May 24, 2021 at 8:13

3 I would also choose Steve's: ) – Fedor Petrov May 24, 2021 at 8:44