How to lie with Statistics: Polish presidential election 2025

true

27/6/2025

Introduction

An explanation of the entirely undeservedly high-profile analysis by Dr. Eng. Kontek, which purported to scientifically prove mass fraud committed during the 2025 presidential election in Poland. For reasons unknown, Mr. Kontek was so convincing that the Prosecutor General, Mr. Bodnar, took his claims with full seriousness:

The election protests filed by Krzysztof Kontek and Joanna Staniszkis were forwarded by the Supreme Court to the Prosecutor General on June 23, 2025, for review and response, with a two-day deadline set for submitting a position, i.e., by the end of June 25, 2025, in accordance with the rules for calculating deadlines set out in the Code of Civil Procedure.

The protest authors attached materials indicating possible irregularities in the vote-counting process. Their allegations were based on a scientifically verifiable method, which led the Prosecutor General to conclude that there was a high probability of actual irregularities in the work of the indicated electoral commissions.

In this short text, we will explain why Mr. Bodnar has embarrassed himself and the offices he holds.

Kontek's Method

Conceptually, the method is very simple:

1. Kontek assumes that within small geographical areas, voting preferences are similar (scientifically homogeneous), which means that support for candidates should be roughly the same across all precincts in each area. He identifies 2,200 such areas. On average, each area contains about 15 precincts and 10,000 voters who cast their ballots (20 million / 2,200).
2. For each area, he calculates the median and the variability of support for Karol Nawrocki. Variability here means how much the results differ between precincts. To measure this variability, he uses something called the median absolute deviation (MAD).
3. He then calculates the following measure:

$$k = \frac{x - me}{mad}$$

where x is the support for KN, me is the median support, and mad is the measure of variability known as MAD. The formula may look intimidating, but if $(x - me)$ is the percentage difference between the support in a precinct and the median for a given group, then dividing it by mad rescales the result from a percentage into a multiple of MAD. So instead of a difference of, say, 4.34%, the result is 2.3 — meaning the deviation from the median is 2.3 MADs.

He then arbitrarily chooses a threshold value he considers acceptable. Values with an absolute magnitude greater than 3 are treated as anomalies. Kontek uses either $k = 2$ or $k = 3$ as his "acceptable" thresholds.

Out of around 32,000 precinct-level election commissions, he thus identified 1,400 anomalies using $k = 3$, or 5,000 anomalies using $k = 2$

Why is this analysis flawed?

Item (assumption) 1 — that small geographical areas must exhibit the same level of support — is pulled out of thin air. He simply assumes it, but there's no reason it must be true. His entire "model" is based on this, and **any model is only as credible as the assumptions it's built on**.

It's worth noting that you don't need any knowledge of statistics to recognize that his work is worthless garbage — just plain common sense is enough. And this is what Justice Minister Bodnar took to the Supreme Court? Honestly, it's embarrassing.

Items 2–4 are just irrelevant math. Of course his method produces "anomalies" — it has to. But whether these are signs of fraud is a completely different question. Kontek simply used an old trick: mathematically proving something, then pretending that this something has to do with something else, despite never actually demonstrating any relationship between the two.

Example

I had not time to recreate the exact grouping of precincts the way Kontek did, but the grouping itself shouldn't matter. Which specific precincts form a group is irrelevant, as long as they are geographically adjacent and there are around 15 of them (otherwise, the whole method is fundamentally flawed).

So, for example, the municipality of Sopot has 25 voting precincts. (Kontek grouped by 15, but the more precincts in a group, the harder it is to detect "anomalies" — so my example is more than valid. The more precincts, the more difficult it is to find "anomalies"!)

After applying Kontek's procedure, we get: median = 26.99% support for KN (obviously, the elite votes for the right candidate :-)) mad = 2.95%.

For two precincts, $k > 3$ (top two rows in the table), that's where they cheated:

siedziba	lgw2	gangus&sutner	median	mad	k
Sanatorium Uzdrowiskowe "Leśnik", ul. 23 Marca 105, 81-820 Sopot	163	49.69	26.99	2.95	7.69
SP ZOZ Sanatorium Uzdrowiskowe MSWiA w Sopocie, ul. Bitwy pod Płowcami 63/65, 81-731 Sopot	264	38.64	26.99	2.95	3.94
Pomorskie Centrum Reumatologiczne im. dr Jadwigi Titz-Kosko w Sopocie Sp. z o. o., ul. 23 Marca 93, 81-820 Sopot	34	32.35	26.99	2.95	1.82
Przedszkole z Oddziałami Integracyjnymi Nr 12, ul. Oskara Kolberga 8, 81-881 Sopot	990	31.62	26.99	2.95	1.57
Zespół Szkół Technicznych, ul. Wejherowska 1, 81-814 Sopot	644	31.06	26.99	2.95	1.38
Szkoła Podstawowa z Oddziałami Integracyjnymi Nr 1 (wejście od ul. Bocznej), ul. Armii Krajowej 50/54, 81-843 Sopot	1230	30.81	26.99	2.95	1.29
Centrum Kształcenia Ustawicznego im. Bohaterów Wybrzeża, ul. Tadeusza Kościuszki 22 - 24, 81-704 Sopot	921	30.73	26.99	2.95	1.27
Szkoła Podstawowa z Oddziałami Sportowymi Nr 7 im. Tadeusza Kościuszki, ul. Jana Jerzego Haffnera 55, 81-715 Sopot	839	30.04	26.99	2.95	1.03
II Liceum Ogólnokształcące im. Bolesława Chrobrego, ul. Aleja Niepodległości 751, 81-838 Sopot	1173	29.41	26.99	2.95	0.82
Sopockie Ognisko Plastyczne, ul. Księżycowa 3b, 81-821 Sopot	945	28.89	26.99	2.95	0.64
Przedszkole Nr 2 im. Jana Brzechwy, ul. 23 Marca 90, 81-820 Sopot	1190	28.49	26.99	2.95	0.51
Dom Pomocy Społecznej, ul. Adama Mickiewicza 49, 81-866 Sopot	904	27.10	26.99	2.95	0.04
Państwowa Galeria Sztuki, pl. Plac Zdrojowy 2, 81-720 Sopot	904	26.99	26.99	2.95	0.00
Zespół Szkół Specjalnych Nr 5 im. Marii Grzegorzewskiej, ul. Kazimierza Wielkiego 14, 81-780 Sopot	1198	26.88	26.99	2.95	-0.04
Szkoła Podstawowa z Oddziałami Integracyjnymi Nr 9 im. Gen. Władysława Sikorskiego, ul. Oskara Kolberga 15, 81-881 Sopot	1073	26.65	26.99	2.95	-0.11
Pomorskie Centrum Reumatologiczne im. dr Jadwigi Titz-Kosko Zakład Rehabilitacji Leczniczej, pl. Plac Zdrojowy 3, 81-720 Sopot	1089	26.45	26.99	2.95	-0.18

siedziba	lgw2	gangus&sutener	median	mad	k
Spółdzielnia Mieszkaniowa im. J. I. Kraszewskiego, ul. Józefa Kraszewskiego 31, 81-815 Sopot	1249	26.34	26.99	2.95	-0.22
Fundacja Nasz Przyjazny Dom, ul. Władysława IV 1c, 81-703 Sopot	663	25.94	26.99	2.95	-0.36
Szkoła Podstawowa z Oddziałami Integracyjnymi Nr 8 im. Jana Matejki, ul. Józefa Golca 3, 81-743 Sopot	1261	25.93	26.99	2.95	-0.36
Niepubliczna Szkoła Podstawowa Las Academy, ul. Kujawska 50/52, 81-862 Sopot	1049	25.74	26.99	2.95	-0.42
Muzeum Sopotu, ul. Księcia Józefa Poniatowskiego 8, 81-724 Sopot	1217	25.72	26.99	2.95	-0.43
Pomorskie Centrum Reumatologiczne im. dr Jadwigi Titz-Kosko w Sopocie Sp. z o.o., ul. Grunwaldzka 1-3, 81-759 Sopot	24	25.00	26.99	2.95	-0.67
Szkoła Podstawowa z Oddziałami Integracyjnymi Nr 1 (wejście od ul. Armii Krajowej i ul. Bocznej), ul. Armii Krajowej 50/54, 81-843 Sopot	955	24.92	26.99	2.95	-0.70
Uniwersytet Gdański, ul. Armii Krajowej 119/121, 81-824 Sopot	1170	23.33	26.99	2.95	-1.24
Przedszkole Nr 8, ul. Jana z Kolna 3, 81-746 Sopot	1888	23.31	26.99	2.95	-1.25

The columns `lgw2` and `gangus&sutener` contain, respectively, the number of valid votes and the percentage of support for KN.

Conclusions?

1. The rural rabble came to Zoppot for treatment and lowered Rafał Trzaskowski support (just kidding). (explanation for foreigners: those 2 anomalities are locate in rehabilitation centers, populated for obvious reasons with people not from Sopot)
2. In 99% of those 1,400 or 5,000 “anomalous” precincts identified by Kontek, the situation is exactly the same as in Sopot.