Australian Bureau of Statistics

Senate Election 2025: impact of assumed exception rate

Australian Bureau of Statistics advice to Australian Electoral Commission

July 2025

Contents

Introduction

The Australian Electoral Commission (AEC) has asked the Australian Bureau of Statistics (ABS) to provide advice on interpreting the results of a statistical assurance process presenting an estimated exception rate for the Senate ballots, calculated after an examination of a sample of ballots. The AEC has previously asked the ABS to advise on methods for calculating the probability of an election result being altered given a specified exception rate, and this report presents updated advice based on results from the 2025 Senate election. The AEC has advised that there have been no changes to the Senate ballot counting rules since the previous advice provided by the ABS, and the AEC has provided the necessary results from the ballot assurance process for the 2025 Senate election.

Scope of the report

This report will:

1. Briefly present context for the ABS advice.
2. Present a formula for estimating the probability of an election result being altered by exceptions and describe how to apply the formula.

A derivation for the formula in Section 2 will be included in an appendix. Table 1 in Section 2 contains an assessment of the results for New South Wales, Victoria and Western Australia.

The analysis has been requested for New South Wales, Victoria and Western Australia by the AEC because the relevant exclusion/election margin was less than twice the maximum estimated exception rate.

Section 1: Context

The AEC has commissioned an assurance of the Senate ballot capture process. Ballots were selected through a random sampling process and examined to determine the proportion of ballots that were flagged as incorrectly interpreted, for example by being assigned to the wrong candidate, having the formal preference sequence broken too early or too late, or by being incorrectly classified as formal or informal. The results of the statistical assurance process are used to calculate an overall exception rate.

The statistical assurance processes have found no indication of any systematic effect among the exceptions. That is, there was no evidence of any candidates or parties being systematically advantaged or disadvantaged by the exceptions.

The ABS has previously provided advice to the AEC on the interpretation of the results. The current report presents updated advice on the interpretation of the results for the 2025 Senate election. There have been no changes to the methodology used to calculate the results in the previous advice provided in 2023

Section 2: Estimating the probability of a different result, given an assumed exception rate.

This section presents a step-by-step method for calculating the probability that an election result is altered by exceptions, given an assumed exception rate and the difference in vote totals between two candidates of interest. It is important to note that the calculations described in this section rely on the assumption, as supported by the findings of the statistical assurance processes, that there are no systematic effects among the exceptions.

The conceptual approach is to calculate the expected number of exceptions within an electorate and then to imagine reassigning the votes to the correct candidates. Given the assumption of no systematic effect, this can be modelling by reassigning the exceptions randomly to different candidates. On average, we would expect a loss of votes from a candidate to be approximately cancelled out by a gain in votes from another candidate. However, in some cases, simply by random chance, it may be possible for a higher proportion of votes to be reassigned to one candidate. The aim of the following calculations is to quantify the chance of this happening to an extent large enough to affect the outcome of an election.

The following step-by-step calculation gives an estimate of the probability of the exceptions causing a change in the outcome of an election:

1. Calculate d = v1 - v2, where v1 is the vote count for the candidate with the highest number of votes and v2 is the vote count for the candidate with the second highest number of votes
2. Calculate x = d(1 - r), where r is the assumed exception rate
3. Calculate the z-score: z = x/(4Nr/(K-1))^0.5, where N is the total number of votes and K is the number of candidates

Calculate the probability Pr(Z>z) for the z-score from Step 3, by looking up standard normal z-tables, or using functions such as “pnorm” in the R statistical software or “norm.s.dist ” in Microsoft Excel. This is the probability that the exceptions cause a change in the outcome between the two candidates.

Example (2025 Senate Election in the state of NSW)

This worked example uses numbers from the 2025 Federal election for the NSW Senate. The figures used in this section have been provided to the ABS by the AEC, and some supplementary information was taken from the distribution of preferences information listed on the AEC website: SenateDistOfPrefsStatisticsReport.

Step 1: The “last difference” (that is, the difference between the number of votes for the last elected candidate and the next, not elected, candidate) was d = 17,326.

Step 2: The estimated maximum proportion of exceptions in NSW was r=0.011. The last difference is multiplied by (1-0.011) to give x = d(1 - r) = 17,135.

Step 3: The final two candidates (that is, the last elected candidate and the next, not elected, candidate) together received N=1,252,638 votes in total. Using the formula above, the z-score for this case is:

$$z = \frac{x}{(4Nr/(K-1))^{0.5}} = \frac{17135}{27558^{0.5}} = 72.99$$

This z-score can be used to calculate a probability that the difference between the votes gained by the last elected candidate and the votes gained by the next candidate is more than 17,135. An online calculator or another tool can show that probability for this example is extremely small.

By way of illustration, the probability that the difference in votes gained is more than 500 is 1.76%.

The probability that the difference in votes gained is more than 1,000 is 0.00126%.

The probability that the difference in votes gained is more than 17,135 is incalculably small.

It should be noted that these calculations rely on a number of assumptions. In particular, these calculations assume that there is no systematic effect among the exceptions, and correcting the exceptions will only results in votes being reassigned to one of the viable remaining candidates.

Summary of assessment of results for NSW, VIC and WA

Table 1 shows the results of the assessment using the formula described above, to calculate the probability of a different result between the last elected candidate and the next, not elected, candidate due to ballot exceptions. Here, the "Votes received by last two candidates" is the sum of the number of votes received by the last elected candidate and the next, not elected, candidate. A value of "~0" for the probability of a different result indicates that the probability is incalculably small.

Table 1: Assessment of results for New South Wales, Victoria and Western Australia

State/Territory	Maximum estimated exception rate	Last difference (d)	Votes received by last two candidates (N)	z-score	Probability of different result
NSW	1.10%	17,326	1,252,638	72.99	~0
VIC	1.00%	32,041	985,947	159.73	~0
WA	1.03%	8,397	406,309	64.23	~0

Appendix: derivation of the formula for calculating the probability of a different result

This appendix outlines the derivation of the formula used to calculate the probability of obtaining a different result, given the estimated exception rate. This has been used in the calculations for Table 1 and the spreadsheet supplied to the AEC by the ABS.

Our goal is to calculate the probability that the election result determined by the observed ballots would have been different if all the ballots were correct. This is denoted by:

$$\Pr\left(V_k^{(true)} - V_j^{(true)} > 0 \middle| V_1^{(obs)}, \dots, V_K^{(obs)}\right)$$

where $V_k^{(true)}$ is the true number of votes cast for candidate *k* (that is, the number of votes that would have been cast for candidate *k* if all the ballots were correct), $V_j^{(true)}$ is the true number of votes cast for candidate *j*, and $V_1^{(obs)}, \dots, V_K^{(obs)}$ are the number of observed votes (that is, the number of votes counted in the ballot process) for each of the candidates in the election, from candidate 1 to candidate K.

The observed votes can be written as follows

$$V_k^{(obs)} = \sum_j V_{j|k}$$

where $V_{j|k}$ is number of votes for candidate j that ended up being assigned to candidate k (either by exception when $j \neq k$ or by assigning correctly when $j = k$)

The true number of votes is given by

$$V_k^{(true)} = \sum_j V_{k|j}$$

The model for exceptions is used to derive a multinomial distribution for $\{V_{1|k}, \dots, V_{K|k}\}$:

$$\{V_{k|1}, \dots, V_{k|K}\} \sim \text{Multinomial}\left(V_k^{(true)}, s_{1k}, \dots, s_{Kk}\right) \quad k = 1, \dots, K$$

$$\text{where} \quad s_{jk} = \begin{Bmatrix} r\dfrac{p_j}{1 - p_k} & k \neq j \\ 1 - r & k = j \end{Bmatrix}$$

If we set $p_k = \frac{1}{K}$ for an assumed even distribution of exceptions, then

$$s_{jk} = \begin{Bmatrix} \dfrac{r}{K - 1} & k \neq j \\ 1 - r & k = j \end{Bmatrix}$$

We can also use the relationship between the Poisson and multinomial distributions to write

$$V_{k|j} \sim \text{Poisson}\left(\alpha_k s_{jk}\right)$$

Then, conditional on $\sum_j V_{k|j} = V_k^{(true)}$, we recover the multinomial distribution. However, we do not actually know what $V_k^{(true)}$ is; instead, we know:

$$V_k^{(obs)} = \sum_j V_{j|k}$$

So, we condition on this to obtain a different set of multinomial distributions:

$$\{V_{1|k}, \dots, V_{K|k}\} \sim \text{Multinomial}\left(V_k^{(obs)}, q_{k1}, \dots, q_{kK}\right) \quad \text{where} \quad q_{kj} = \frac{\alpha_j s_{kj}}{\sum_l \alpha_l s_{kl}} \quad j, k = 1, \dots, K$$

If the value for $\alpha_k = \alpha_j$ is assumed to be equal for different candidates, and the exception parameters are $p_k = \frac{1}{K}$ then we obtain the same probabilities for q_{kj} as for s_{kj}:

$$\alpha_1 = \cdots = \alpha_K = \alpha \qquad p_1 = \cdots = p_K = \frac{1}{K} \quad \Rightarrow \quad q_{kj} = s_{kj} = \begin{Bmatrix} \dfrac{r}{K - 1} & k \neq j \\ 1 - r & k = j \end{Bmatrix}$$

Given the above model, we express the difference in true votes in terms of components:

$$V_k^{(true)} - V_j^{(true)} = \sum_l \left(V_{k|l} - V_{j|l}\right)$$

The expected value for this is:

$$E\left(V_k^{(true)} - V_j^{(true)}\right) = V_k^{(obs)} q_{kk} + \sum_{l \neq k} V_l^{(obs)} q_{lk} - V_j^{(obs)} q_{jj} - \sum_{m \neq j} V_m^{(obs)} q_{mj}$$

Calculating the variance is more complex, as there is some correlation between terms from the same multinomial distribution. However, we can assume independence across the multinomial distributions:

$$Var\left(V_k^{(true)} - V_j^{(true)}\right) = \sum_l Var\left(V_{k|l} - V_{j|l}\right)$$

$$= \sum_l Var\left(V_{k|l}\right) + Var\left(V_{j|l}\right) - 2Cov\left(V_{k|l}, V_{j|l}\right)$$

$$= \sum_l \left[V_l^{(obs)} q_{lk}(1 - q_{lk}) + V_l^{(obs)} q_{lj}(1 - q_{lj}) - 2\left(-V_l^{(obs)} q_{lk} q_{lj}\right)\right]$$

$$= \sum_l V_l^{(obs)} \left[q_{lk} + q_{lj} - \left(q_{lk} - q_{lj}\right)^2\right]$$

Simplifying this gives:

$$E\left(V_k^{(true)} - V_j^{(true)}\right) = \left(V_k^{(obs)} - V_j^{(obs)}\right)(1 - r) + \frac{r}{K-1}\left(\sum_{l \neq k} V_l^{(obs)} - \sum_{m \neq j} V_m^{(obs)}\right)$$

$$= \left(V_k^{(obs)} - V_j^{(obs)}\right)(1 - r) + \frac{r}{K-1}\left(\sum_l V_l^{(obs)} - V_k^{(obs)} - \sum_m V_m^{(obs)} + V_j^{(obs)}\right)$$

$$= \left(V_k^{(obs)} - V_j^{(obs)}\right)(1 - r) + \frac{r}{K-1}\left(V_j^{(obs)} - V_k^{(obs)}\right)$$

$$= \left(V_k^{(obs)} - V_j^{(obs)}\right)\left(1 - r - \frac{r}{K-1}\right)$$

$$= \left(V_k^{(obs)} - V_j^{(obs)}\right)\frac{K(1-r) - 1}{K-1}$$

$$= \hat{\delta}_{jk}$$

$$\approx V_k^{(obs)} - V_j^{(obs)} \quad (\text{assuming } \; r \approx 0)$$

And so,

$$Var\left(V_k^{(true)} - V_j^{(true)}\right)$$

$$= V_k^{(obs)}\left[q_{kk} + q_{kj} - \left(q_{kk} - q_{kj}\right)^2\right] + V_j^{(obs)}\left[q_{jk} + q_{jj} - \left(q_{jk} - q_{jj}\right)^2\right] + \sum_{l \neq j,k} V_l^{(obs)}\left[q_{lk} + q_{lj} - \left(q_{lk} - q_{lj}\right)^2\right]$$

$$= V_k^{(obs)}\left[1 - r + \frac{r}{K-1} - \left(1 - r - \frac{r}{K-1}\right)^2\right] + V_j^{(obs)}\left[\frac{r}{K-1} + 1 - r - \left(\frac{r}{K-1} - 1 + r\right)^2\right] + \sum_{l \neq j,k} V_l^{(obs)}\left[\frac{r}{K-1} + \frac{r}{K-1} - \left(\frac{r}{K-1} - \frac{r}{K-1}\right)^2\right]$$

$$= V_k^{(obs)}\left[1 - r + \frac{r}{K-1} - \left(1 - r - \frac{r}{K-1}\right)^2\right] + V_j^{(obs)}\left[\frac{r}{K-1} + 1 - r - \left(\frac{r}{K-1} - 1 + r\right)^2\right] + \sum_{l \neq j,k} V_l^{(obs)}\left[\frac{r}{K-1} + \frac{r}{K-1} - \left(\frac{r}{K-1} - \frac{r}{K-1}\right)^2\right]$$

$$= \left(V_k^{(obs)} + V_j^{(obs)}\right)\left[r(1 - r)\frac{K+1}{K-1} + \frac{r}{K-1}\left(1 - \frac{r}{K-1}\right)\right] + 2\frac{r}{K-1}\sum_{l \neq j,k} V_l^{(obs)}$$

The underlying z-score is therefore:

$$z = \frac{0 - \hat{\delta}_{jk}}{\sqrt{\left(V_k^{(obs)} + V_j^{(obs)}\right)\left[r(1-r)\frac{K+1}{K-1} + \frac{r}{K-1}\left(1 - \frac{r}{K-1}\right)\right] + 2\frac{r}{K-1}\sum_{l \neq j,k} V_l^{(obs)}}}$$

If we substitute $V_l^{(obs)} = \frac{N}{K}$ (that is, replace the observed votes with the expected votes under the model) then we obtain

$$Var\left(V_k^{(true)} - V_j^{(true)}\right) \approx \left(\frac{N}{K} + \frac{N}{K}\right)\left[r(1-r)\frac{K+1}{K-1} + \frac{r}{K-1}\left(1 - \frac{r}{K-1}\right)\right] + 2\frac{r}{K-1}\sum_{l \neq j,k} \frac{N}{K}$$

$$= 2\frac{N}{K}r\left[(1-r)\frac{K+1}{K-1} + \frac{1}{K-1}\left(1 - \frac{r}{K-1}\right) + \frac{K-2}{K-1}\right]$$

$$= 4\frac{N}{K}r\left(\frac{K}{K-1}\right)\left(1 - \frac{r}{2}\frac{K}{K-1}\right)$$

$$= \frac{4Nr}{K-1}\left(\frac{K\left(1 - \frac{r}{2}\right) - 1}{K-1}\right)$$

$$\approx \frac{4Nr}{K-1}$$

Australian
Bureau of
Statistics

www.abs.gov.au