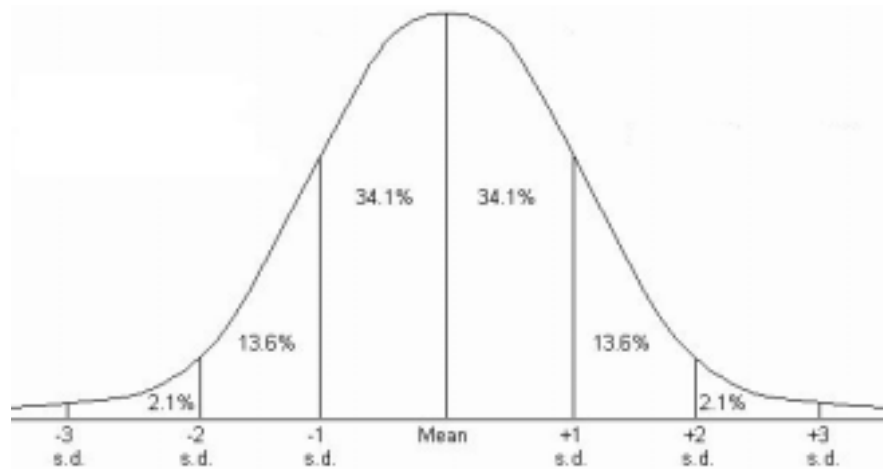# Departure From Expectation

E Bruce Brooks
University of Massachusetts at Amherst
WSWG Note 56rev (10 Feb 1996)

We expect one thing, we get another. You had 10 students in your class last year; but this year you had 15. There were 112 incidents of car theft last year; this year, 121. Is this significant? Such questions come up all the time. They have a simple solution.

Normal Expectation

If occurrences diverge randomly from an expected value, the pattern of those departures is given by the normal distribution, whose curve looks like this:

All that means is that if our variable is varying randomly around a norm, 68·2% of all occurrences will fall within ±1 standard deviations (sd), and 95·4% within ±2 sd. If the matter is important, we want instead a 99% assurance that a given departure from expectation is meaningful, not random. That level is defined by ±2·54 sd.

For a given pair of E (expected) and A (actual) numbers, how do we figure the sd? The usual approximation formula[1] is:

$$(A - E) / \sqrt{E}$$

Which is simple enough. But to make it self-interpreting, so that the formula will give 1·00 when the 99% level of assurance is reached, we multiply by 0·39.[2]

---

[1]For the basic formula, see Paul G Hoel **Elementary Statistics** (2ed Wiley 1966) 103-106.

[2]Multiplying by 0·39 is the same as dividing by 2·54, the number of standard deviations.

The absolute value of (A – E) needs to be used in calculating, and any minus sign should be appended to the final result[3]. A further factor with human data is that once something unusual occurs (a rare word in a text) it is likely to repeat (a copycat crime). To compensate for this, we take the square root of the result. The final formula is:

$$S = \sqrt{\ [(0{\cdot}39)(A-E) / \sqrt{E}\ ]}$$

This can be done in seconds on a hand calculator, as long as it has a square root key.

Practical Applications

1. Your class had 10 students last year; the expectation (E) is for 10 this year too. You actually (A) had 15. The significance (S) of this works out to

$$\sqrt{[(0{\cdot}39)(15-10) / \sqrt{10}]} = \sqrt{(0{\cdot}39)(5) / (3{\cdot}16)} = \sqrt{0{\cdot}61} = +0{\cdot}78,\ \textbf{not}\ \text{significant}$$

2. Over the 242 years covered by the Lǔ chronicle Chūn/Chyōu (CC), there are 524 military events, or 2·17 per average year. For the 18-year reign of Lǔ Hwán-gūng, we thus expect 39·06 military events; instead there are actually 16. The significance is:

$$\sqrt{\ [(0{\cdot}39)(16-39{\cdot}06) / \sqrt{39{\cdot}06}]} = \sqrt{(0{\cdot}39)(-23{\cdot}06) / (6{\cdot}25)} = \sqrt{-1{\cdot}44} = -1{\cdot}20,\ \textbf{significant}$$

3. The average number of CC diplomatic events is 2·73 per year; for Hwán-gūng we expect 49·14 and actually get 58. The significance of this is:

$$\sqrt{[(0{\cdot}39)(58-49{\cdot}14) / \sqrt{49{\cdot}14}]} = \sqrt{(0{\cdot}39)(8{\cdot}86) / (7{\cdot}01)} = \sqrt{0{\cdot}49} = +0{\cdot}70,\ \textbf{not}\ \text{significant}$$

but this plus the preceding result suggests that the dip in military events may have had something to do with diplomatic efforts by Hwán-gūng, who went often to other states.

4. ευσεβεια "religion" occurs 15× in the New Testament (138,019 words), never in Paul, but 10× (A) in the post-Pauline Pastoral Epistles (9,488 words, so E = 1·03).

$$\sqrt{\ [(0{\cdot}39)(10-1{\cdot}03) / \sqrt{1{\cdot}03}]} = \sqrt{(0{\cdot}39)(8{\cdot}97) / (1{\cdot}01)} = \sqrt{3{\cdot}46} = +1{\cdot}86,\ \textbf{significant}$$

This reflects the conventionalizing of Christian belief in the years after Paul.

5. The following are the actual crime statistics for a certain city in 1994 and 1995,[4] and their calculated significance:

	1994 (E)	1995 (A)	S	Interpretation
Murder	0	1	+0·62	not significant
Sexual Assault	55	47	−0·65	not significant
Assault/Battery	295	308	+0·54	not significant
Breaking/Entering	91	109	+0·86	**not yet** significant
Car Theft	112	121	+0·58	not significant
Vandalism	571	520	−0·91	**not yet** significant

All this should be watched by an alert Police Chief. There is no present need to shift staff from one detail to another, but notice: profitless crimes (vandalism) are down; profitable crimes (breaking/entering) are up. Creeping urbanization?

S, as here presented, is for situations adequately described by one A and one E. Any real-life Chief would of course be tracking crimes over more than one year.

---

[3]A more elaborate way to say this is to multiply the final result by (A - E) / | A - E |.

[4]From the Daily Hampshire Gazette, 21 Feb 1996. As a matter of retrospective compassion, I do not here discuss the paper's own interpretation of these figures.