2. Z-SCORES AND THE EMPIRICAL RULE

Eg 1: The earnings per share (EPS) for FedEx in the second quarter were $2.66, failing to meet analysts' expectations of $2.70 (consensus forecast). The earnings surprise for FedEx was $2.66 - 2.70 = -0.04. Is this bad compared to the earnings surprise of five other NYSE companies observed in a similar time period?
<table>
<thead>
<tr>
<th>Company</th>
<th>Surprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kroger</td>
<td>.04</td>
</tr>
<tr>
<td>Red Hat</td>
<td>.04</td>
</tr>
<tr>
<td>Rite Aid</td>
<td>.02</td>
</tr>
<tr>
<td>Pier 1</td>
<td>.00</td>
</tr>
<tr>
<td>Oracle</td>
<td>−.08</td>
</tr>
</tbody>
</table>
**Preliminary Calculations:** The other earnings surprise values are .04, .04, .02, 0, −.08. The sample mean is

\[ \bar{x} = (1/5)(.04 + .04 + .02 + 0 − .08) = .02 / 5 = .004. \]

The sample variance is \( s^2 = (1/4)[(.04 − .004)^2 + (.04 − .004)^2 \]
\[+ (.02 − .004)^2 + (0 − .004)^2 + (−.08 − .004)^2 ] \]
\[= .00992 / 4 = .00248. \]

The sample standard deviation is \( s = \sqrt{.00248} = .0498. \)

How does this help us put the surprise for FedEx (−.04) into context?

To answer this, we need to define the z-score.
Suppose $x$ is a value, or “score”. We want to compare $x$ to a data set (or population) with mean $\bar{x}$ (or $\mu$) and standard deviation $s$ (or $\sigma$). $x$ may or may not be a part of the data set.

Define the $z$-score for $x$ as $z = \frac{x - \bar{x}}{s}$ or $z = \frac{x - \mu}{\sigma}$.

• The $z$-score tells us how many standard deviations $x$ is from the mean.

• The further $z$ is from zero, the more “atypical” $x$ is, relative to the given data set. In fact, the “empirical rule” states that for roughly bell-shaped distributions:
  about 68% of the data values will have $z$-scores between $\pm 1$,
  about 95% between $\pm 2$, and
  about 99.7% (i.e., almost all) between $\pm 3$. 
Thus, for example, roughly 95% of the data should lie within two standard deviations of the mean, that is, in the interval $(\mu-2\sigma, \mu+2\sigma)$.

The empirical rule is only exactly correct for normal distributions (using true population means and standard deviations), but it often provides a useful approximation, even for non-normal distributions.

It is worth memorizing the empirical rule.
Eg 1, contd.: FedEx had an earnings surprise of \(-.04\). Relative to the other 5 firms, having \(\bar{x} = .004\), \(s = .0498\), the \(z\)-score for FedEx is

\[
z = \frac{x - \bar{x}}{s} = \frac{-0.04 - 0.004}{0.0498} = -0.88 .
\]

The earnings surprise for FedEx was less than one standard deviation below the mean, and therefore quite typical.

The empirical rule helps us to put variation in context. But variation cannot be completely tamed.

Overadjustment Increases Variation: Deming’s “Funnel” Experiment. (In-Class Demo).