# BASIC STATISTICS FOR R & D

**by Jay Apt**

Many R & D projects base their conclusions on quantitative experimental data taken as part of the project. However, few reports show any concern for taking *a lot* of data at each data point they present, hence make their conclusions shaky at best. Suppose a project sets out to determine whether black rockets go higher than white rockets of the same design. If the experimenter builds the models, launches each one, and reports that the black rocket was tracked to 330 meters, and the white one to 320 meters, most of us would feel that a conclusion that black models do indeed fly higher would be unjustified. The basis for this intuitive feeling is our experience that engine thrust is not constant from one engine to the next, and that tracking really isn't very precise. We'd have been much more confident of the results if he had launched the model hundreds of times and reported that the difference in the average altitudes was greater than could be accounted for by tracking errors.

This example illustrates two of the many types of uncontrollable external variables whose fluctuations can cause a single piece of data to be far from an average obtained by making a large numbers of experimental runs. If these fluctuations are as likely to be negative as positive, they are called *random* fluctuations, since they introduce a random error in any single piece of data that cannot be predicted. For example, the variation in engine thrust discussed previously introduces a random error in altitude. Or, if you perform a Gregorek-type strobe B/G analysis (*Model Rocketry*, November, 1971 and NARTS TER 4), the speed at which you throw the glider can also be considered as fairly random. It should be clear that for this type of fluctuation-where the error from the average is as likely to be positive as negative-that the more data you take, the closer your average value is likely to be to the true result.

If you are to convince anyone that the conclusion you draw from your data is valid, it is very important that you make some estimate of the error due to these random fluctuations. For even if you launch your model in an altitude experiment 20 times, you can't be sure that you have averaged out the engine and tracking fluctuations. What we want to be able to do is report data with an error estimate-330 ± 40 meters for example-which says that we are reasonably confident that the average altitude we would get if we launched the model a hundred thousand times would be somewhere between 290 and 370 meters.

Suppose that the day before we do our experiment, the ghost of Dr. Goddard sneaks out to the launch field and *does* launch our model a hundred thousand times, with closed tracks each flight. If he uses, say, a B engine with a rated total impulse of 5.00 newton-seconds, he may get some engines with very little thrust, and some with as much as 10 or 15 nt-sec, all because the engine-making machines aren't perfect. So he'll have some flights that barely get off the pad, while others will make him swear that our model will win design efficiency at the next contest. A graph of his altitudes would then be like that shown in Figure 1. (This bell-shaped curve is familiar because many types of fluctuations follow the same principle as engine variations: the likelihood of any one piece of data varying from the average a small amount is high, and although it doesn't happen very often, data sometimes varies a large amount from the average.)

When we go out to the field the next day, we can't take nearly as many data points as Dr. Goddard did, so our data might look like that shown in Figure 2. In general, the average of our data won't be the same as his, since we probably haven't averaged out all the fluctuations in engines, tracking, etc. What must now do is mathematically analyze our data to make an estimate of how far off the true average our average from the 20 flights is; that is , compute probable error limits.

Even though we have taken only a few data points, they should fall on the same shaped curve as Dr. Goddard's, since any collection of data is just a sample of the data that would be recorded if the experiment were repeated forever. So we can use this bell-shaped curve as a model for our data, and use the mathematical analysis of this type of curve that was developed long ago. What we are interested in is the best possible estimate of how far the average of our data is from the average of the data from the experiment repeated forever. In mathematical terms, this best estimate is called the *standard deviation of* *the mean* ("mean" is just another word for average). This "standard deviation" is a standardized "best estimate" of the difference of our average from the true average. It is standardized in such a way that if you report an altitude of 330 meters with a standard deviation of the mean of ±40 meters you are saying that you are 68% confident that the true average lies within one standard deviation (±40 meters) of 330, and 95% confident that the true average lies somewhere within ±80, or two standard deviations, of 330.

We calculate the average of our data in the normal way, by adding up all the altitudes and dividing by the number of flights, If N stands for the number of flights X(1), X(2), X(3), etc. are the altitudes of the flights, and $\overline{X}$ is the average, the formula is* :

$$\overline{X}=\frac{1}{N}\sum_{I=1}^{N} X(I)$$

Having once calculated this average, we can get the standard deviation of the mean from the following formula where $\sigma_m$ stands for the standard deviation of the mean:

$$\sigma_m=\frac{1}{\sqrt{(N)\,(N-1)}}\sqrt{\sum_{I=1}^{N}\,[X(I)-\overline{X}]^2}$$

This formula tells us to find the difference between each of the data values and their average, then square that difference. We do this for all of our data values, add all the squared results, and take the square root of that sum. We finally divide this result by the square root of N, (the number of data points) times N-1 to get the standard deviation of our mean. The standard deviation of the mean of 20 data values shown in Figure 2 is worked out in Table 1.

It is this standard deviation of the mean that we should quote as the error attached to our average altitude (here it would be 329 ±4 meters). And the formula for $\sigma_m$ tells us several interesting things. First, we see that we can reduce the standard deviation of the mean simply by taking more data, since it will get smaller as N increases, Second, if our data is widely scattered, the quantity $[X(1)-\overline{X}]$ will be large for most of the data values, so $\sigma_m$ will also be large. So we know that if most of the altitudes in an altitude experiment, for example, are not about the same, we should not expect their average to give very precise information. For example, it would be wrong to draw any conclusion about whether the black rocket goes higher than the white rocket based on data of 330 ±20 meters for the black rocket and 320 ±20 meters for the white rocket, no matter how many flights are included in the data. Analyzing you data in this way in the field isn't very difficult, and can often tell you whether you need to take more data to be able to draw a valid conclusion.

The formulas we have just discussed give an estimate for the uncertainty in you data only for random errors of the type previously illustrated. Keep in mind that things like difference in fin finish, launch lugs, calibration of any measuring apparatus, linearity of data sensors used, and personal errors on the part of the experimenter can put the average of the data taken much farther from the true average than is indicated by these formulas. That's why it's important to fully describe the experimental procedures used in you report.

Finally, note that although we have used an altitude experiment as our example in this article, the same principles of taking a lot of data at each data point and reporting the averaged results with standard deviations of the mean attached apply to *all* R & D projects where quantitative data is taken.

*For those unfamiliar with it, the notation

$$\sum_{I=1}^{N} X(I)$$

is a shorthand that tells us to add X(1) + X(2) + X(3) + ... + X(N).

Jay Apt is a former NAR Secretary and Trustee and is currently a Shuttle astronaut.

Reprinted from the December 1972 *Model Rocketeer.*

Table 1

| X(I) | X(I)−$\bar{X}$ | $[X(I)-\bar{X}]^2$ |
|---|---|---|
| 305 | -24 | 576 |
| 305 | -24 | 576 |
| 310 | -19 | 361 |
| 315 | -14 | 196 |
| 320 | - 9 | 81 |
| 320 | - 9 | 81 |
| 320 | - 9 | 81 |
| 325 | - 4 | 16 |
| 325 | - 4 | 16 |
| 325 | - 4 | 16 |
| 330 | 1 | 1 |
| 330 | 1 | 1 |
| 330 | 1 | 1 |
| 335 | 6 | 36 |
| 335 | 6 | 36 |
| 335 | 6 | 36 |
| 340 | 11 | 121 |
| 345 | 16 | 256 |
| 360 | 31 | 963 |
| 370 | 41 | 1681 |

N=20 $\quad \Sigma[X(I)-\bar{X}]^2=5132$

$\sum_{I=1}^{20} X(I)=6580 \quad \sqrt{5132}=71.6$

$\bar{X}=329 \quad \sigma_m = 3.7 \approx 4$

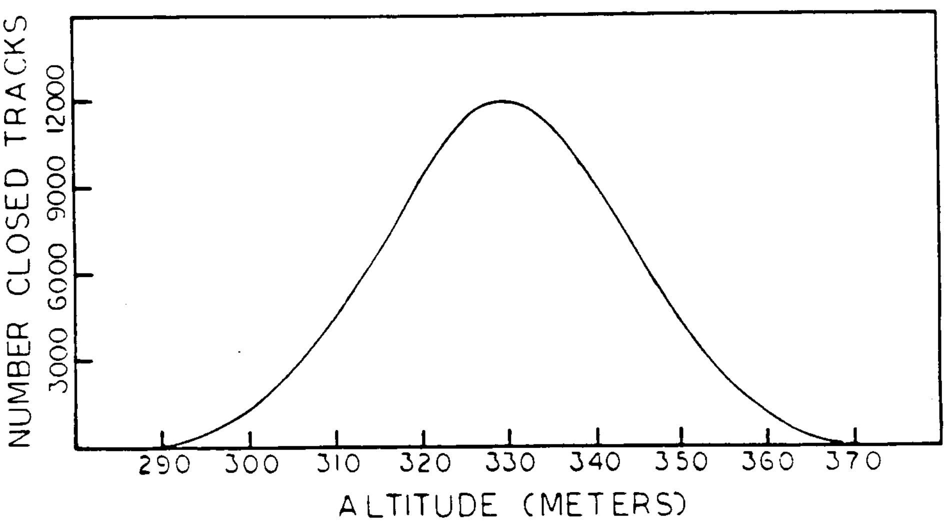


FIGURE 1

FIGURE 2