

PROBLEMS WITH PROBABILITIES

The BGA Safety Team discusses some of the snags with interpreting statistics

The BGA's archive contains over 7,500 accident, incident and technical occurrence reports spanning nearly 50 years of UK and expedition gliding. It's an invaluable safety resource that we consult almost every day. Key details are recorded and categorised in a searchable database: we mainly use this to identify reports to re-read when investigating a particular topic, but it also allows rudimentary 'health checks' to see how different years, clubs, aircraft, etc, compare. We're able to see, for example, that the rate of winch launch accidents has fallen by a factor of five since the introduction of the Safe Winch Launch initiative in 2005; and that collisions between gliders are almost unknown if both aircraft carry FLARM.

Gliding operations, and the associated risks, are quite varied, and we'd be surprised if accident rates and profiles didn't differ between large, flat sites in benign climates and small, ridge sites that experience strong winds and rotor. We might also



expect differences associated with launch method, glider types and fleet profiles, pilot ages, levels of experience, and whether clubs focus upon training, first flights or expeditions.

We'd also suspect that developments in design and technology, ageing airframes, airspace restrictions, or broader changes in society could cause accident rates to change with time; and we'd hope that advances in training, maintenance and safety culture might produce discernible improvements. Happily though, in many cases, we don't have enough accidents to distinguish.

CLUB COMPARISONS

It's not uncommon for a club to ask us how its safety record stands against the broader national picture, and we're happy to oblige. However, such comparisons come with severe caveats, because in many cases even notable differences may not have statistical significance, while apparently encouraging figures could hide a concerning situation.

As an example, a quarter of UK clubs typically log fewer than 1,000 launches per year. At the long-term average accident rate, each club would experience no more than one fatal or serious injury accident every 40 years, and a club could be several times more susceptible and still be likely to have an accident-free decade. On the other hand, we'd expect the 20 or so least active clubs to account between them for several accidents per decade, but the clubs at which these accidents occur are not necessarily more susceptible than the others – it could just be that it's 'their turn'.

FUNNEL PLOTS

A convenient way of representing such data is as a funnel plot [1, 2], commonly used to compare medical outcomes [3]. Figure 1 shows, in proportion to the number of launches, the number of aircraft written off or substantially damaged at different clubs over a 10-year period. These figures are plotted against the number of launches from which the rate was calculated, so that club launch activity increases from left to right. In this example, while the launch figures are typical of BGA clubs, the damage rates come simply from a simulated random distribution for a constant probability. Crucially, the plot also shows lines bounding the random ranges expected given the total launch and accident numbers. If, as in this simulation, accidents occurred completely independently and randomly, and there were no differences between clubs except their launch statistics,

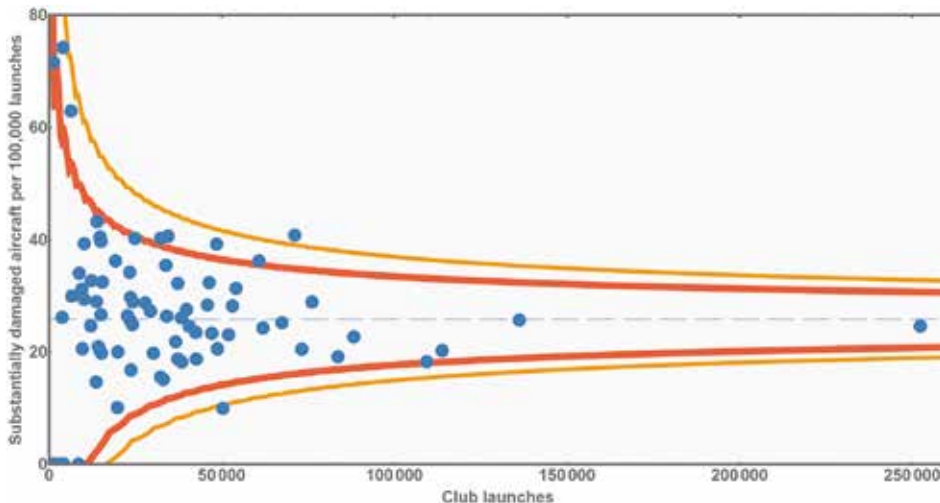


Figure 1: Funnel plot showing random variation in club accident rates simulated for the same mean rate (dashed). Each blue point represents a club's figures for a decade. On average, 10 clubs will lie outside the inner, orange funnel, and two outside the outer, yellow funnel.

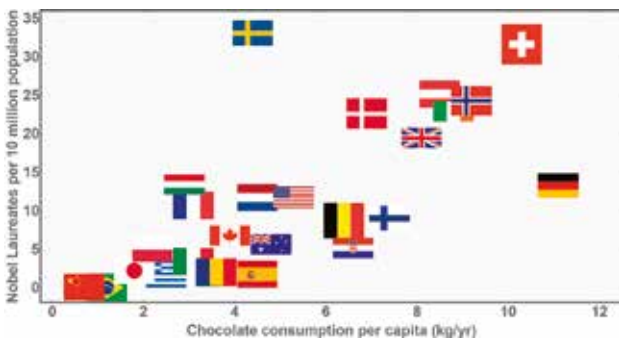


Figure 2: Correlation between chocolate consumption (2019) and Nobel prizes (to date) by country, after [5].

we would expect about 70 clubs to lie within the central region, four clubs in each of the regions above and below that, and two more beyond. In this example, the accident rate for clubs with low launch figures varies from 0-3 times the national average. Real data often look very similar: it could be that there is a variation in accident risk, but the data are consistent with random statistics.

COINCIDENCES

If events occur randomly and independently then the chance of two within any period is the product of their individual probabilities. The coincidence of two rare events therefore suggests that they are not independent, and prompts the quest for a connection. A recent example was the failure, three months apart, of two different Vega undercarriages during the launch ground-run. As we knew of only five other such incidents in 50 years, the chance of two in the same quarter was around 1 in 1,000. Yet, with no change in maintenance or operation and only long-term mechanical degradation, there was no conceivable link between the events.

The explanation is that our records span many rare accidents and incidents, and 200 periods of three months. The chance of two Vega undercarriages failing within three months sometime during our records is around 1 in 5; and we would have been just as intrigued had the coincidence involved any other of the dozens of glider types (though most have more robust undercarriage mechanisms [4]). We should therefore be surprised if there were no such coincidences somewhere in our records.

CORRELATIONS

In the same way that enough sets of random events will eventually reveal coincidences,

then with enough sets of data it is possible to find entertaining correlations. There are, for example only 5,040 permutations of the G7 nations so, amongst thousands of datasets detailing aspects such as population, finance, geography and consumption, there will be many pairs that are similarly ordered.

It is harder to find a perfect match with 26 countries, but the correlation in Figure 2 between chocolate consumption and Nobel prizes is still impressive and has prompted half-hearted searches for a connection [5]. The least tenuous is perhaps that both reflect a country's established affluence, but the pattern is probably just a fluke. We always need to be cautious interpreting such data: correlation does not imply causation.

DATA SHORTCOMINGS

A further problem is that our records are imperfect. Not all accidents are reported, though we like to think that we hear about most; accident reports are incomplete and occasionally contain errors; and we have only poor baseline data about pilot ages, hours flown and so on. Better data might not affect the 'big picture', but it could help us identify new vulnerabilities, and solid evidence has proved invaluable against unjustified regulation.

All this means that our safety records, though a precious and powerful resource, have to be used with care. You can help by ensuring that all accidents and incidents are reported fully and accurately, and are accompanied by information such as photographs and logger traces. We wish we had more data, but uncertainty is welcome when its cause is a low accident rate.

Tim Freearge and the BGA safety team

HAPPILY, WE DON'T HAVE ENOUGH ACCIDENTS TO DISTINGUISH

■ Clubs can obtain printed copies of Safety Briefings and the 2022 accident review [6] from the BGA Office.

[1] B Goldacre, *DIY statistical analysis: experience the thrill of touching real data*, The Guardian (28 Oct 2011) <https://tinyurl.com/flyright2314>

[2] D Spiegelhalter, *Funnel plots for comparing institutional performance*, Statist Med 24, 1185 (2005) <https://tinyurl.com/flyright2315>

[3] Public Health England, *Technical Guide: Funnel Plots* (2017) <https://tinyurl.com/flyright2316>

[4] BGA Technical News Sheet 1-2022 <https://tinyurl.com/flyright2317>

[5] A L Prinz, *Social Sciences & Humanities Open* 2, 100082 (2020) <https://tinyurl.com/flyright2318>

[6] BGA Accident Review 2022 <https://tinyurl.com/flyright2319>

PREVIOUS 'FLY RIGHT' ARTICLES

- The perils of distraction (Apr/May 19)
- Keeping safe in thermals (June/July 19)
- Why it is good to think ahead (Aug/Sep 19)
- The effects of wind gradient (Oct/Nov 19)
- A fun but safe introduction (Dec 19/Jan 20)
- Stop the drop (Feb/Mar 20)
- Avoiding upset (Apr/May 20)
- Backroom boys (June/July 20)
- Cockpit muddle (Aug/Sep 20)
- Safe rotation (Oct/Nov 20)
- Cockpit remedies (Dec 20/Jan 21)
- COVID currency (Feb/Mar 21)
- Eroded margins (Apr/May 21)
- A good lookout (June/July 21)
- Trouble with turbos (Aug/Sept 21)
- 'Hopefully' is not an option (Oct/Nov 21)
- Act when the launch fails (Dec 21/Jan 22)
- Time to solve a knotty problem (Feb/Mar 22)
- RTFM: Read the flight manual (Apr/May 22)
- Startling events (June/July 22)
- Collision risks (Aug/Sep 22)
- Winter hazards (Oct/Nov 22)
- Swiss cheese (Dec 22/Jan 23)
- An expensive mistake (Feb/Mar 23)
- What's changed? (Apr/May 23)
- Aerotow eventualities (June/July 23)