

Using Pareto for Sasquatch Statistics

What would be a good way to determine the ACTUAL number of sasquatch related incidents? In other words both the reported and non-reported incidents. The Italian economist Vilfredo Pareto came up with a way. The following is from Wikipedia:

The Pareto principle (also known as the 80/20 rule, the law of the vital few, or the principle of factor sparsity) states that, for many events, roughly 80% of the effects come from 20% of the causes. Management consultant Joseph M. Juran suggested the principle and named it after Italian economist Vilfredo Pareto, who noted the 80/20 connection while at the University of Lausanne in 1896, as published in his first paper, "Cours d'économie politique". Essentially, Pareto showed that approximately 80% of the land in Italy was owned by 20% of the population; Pareto developed the principle by observing that about 20% of the peapods in his garden contained 80% of the peas.

What this means with regard to the sasquatch is that probably the reported incidents are just 20% of all the incidents.

All we have to do to get the total number of incidents (reported and non-reported) is to apply a little formula ($.20x = \text{Reported Incidents}$). This gives us the number of which 20% represents.

If we use the number of incidents in British Columbia over 100 years (all verified) as our base (379), here's how it works:

$$\begin{aligned} .20x &= 379 \\ x &= 379 / .20 \\ x &= 1,895 \end{aligned}$$

This indicates that the total incidents were 1,895 over that period of time. This can be averaged out to about 19 incidents a year. However, there were far fewer people in BC in the early years, so this would have to be taken into consideration to get the true yearly incidents.

NOTE: If you verify reports and some or many are discarded, they are not in this count because they were not considered valid reports in the first place.

If we just take the 10-year period (2000 to 2009) for which there was 47 incidents,



Vilfredo Pareto (1848–1923). His 80/20 rule is widely used in business. It more or less gives us the “rule” of nature and can be applied to all aspect of human interactions and processes.

then the total number for this period was 235, or 23.5 per year.

The RATIO for the incidents is (235/47) 5:1. In other words, for every reported incident there were 5 non-reported incidents. This ratio would apply to any use of the formula.

The BFRO has 631 published reports for the State of Washington. If we apply the formula, then there were 3,155 reports counting those that were not published. The time frame is about 22 years (the BFRO started in 1995). This indicates that the average incidents per year is 143.

Washington State has 7.288 million people; BC has 4.631 million, so there are far more “human eyes” in Washington than in BC. Also, I think people in the USA are much more “outgoing” than people in BC. Furthermore, Washington has far more access to wilderness regions. Such in BC is very limited. Another factor is that the weather in Washington is somewhat more agreeable than in BC. The further South you go in Washington, the better it becomes. In BC the further North you go the worst it becomes

My grand total for all reported sasquatch incidents in North America (determined in 2003) was 2,557 over 100 years. After applying Pareto's rule we get a total of 12,785, reported and non-reported incidents over this period. Again, if we average this, the result (128 per year) doesn't mean much because incidents depend on people. I originally considered the ratio to be about 8:1

(eight non-reported to every one reported). According to Pareto, it should be 5:1.

Fourteen years have passed since I compiled my sasquatch statistics, so there has definitely been an increase in the total number. Here we can roughly use the average of 128 per year (probably not less) and add 1,792 to the total figure. This equals 14,577. If you want to sort of “reverse engineer” this number, 2,915 were reported incidents, and 11,662 were non-reported.

With the advent of the Internet in the early 1990s, there have been far more reported incidents—mainly due to the ease of reporting using email and the increased number of reporting facilities (websites and researchers). Had the internet been available in 1903, I am sure the total number would be far greater. Furthermore, people are a lot more willing to come forward with information now than in early years. It would be very difficult to factor in these conditions. All you can do is say that the final number would not be less than calculated.

Not many people like numbers, so I will make things simple. If you have collected a number of reports in your region, simply divide the total number by .20 (don't forget the decimal in front). That gives you both the reported and non-reported incidents. Subtract the reported number, and that gives you the non-reported number. Here is an example:

A. Reports Verified:	187
B. Pareto Total (A/.20):	935
C. Non-reported (B-A):	748.

There were 187 Reported Incidents and 748 Non-Reported incidents. These two numbers when added together must give you the total—in this case 935.

Keep in mind that the 748 reports not submitted would have been VERIFIED reports had they been submitted.

Of course, if you think people in your region are far more likely to report incidents, then INCREASE the “.20” (which is 20%). If you think 50% is a better number in your region, then replace the .20 with .50. Just come up with a percentage and change the percentage sign (%) to a decimal (e.g. 37% is .37; 65% is .65 and so forth). I really wouldn't go the other way (reduce the .20) unless you have a very good reason. Pareto kind of set the standard here and he might get upset if we fiddle around with this number.