How many were there when it mattered?

Estimating the sizes of crowds

150 000 people demonstrated in Hong Kong. 10 000 students protested in London against tuition fees. A million people lined the streets for the Royal wedding. Or did they? Do not believe what you are told, say **Ray Watson** and **Paul Yip**. Estimating the size of a crowd is a difficult business – even for those who actually want to get it right.

Counting the size of a crowd has been an issue for at least 2000 years: "and those who ate were about five thousand men, besides women and children", says the Bible (Matthew 14: 21). Claims for the numbers in a crowd have been a part of politics and public relations for a very long time. There is a large amount of variability in crowd estimates, for two reasons: it is difficult to do, and there are strong motivations for getting it wrong!

There are two major types of crowd: those for a public event and those for a demonstration or rally. In either case the crowd may be essentially static, where there is some assembly area which is the focus of the event, or mobile, where progress is part of the process.

Royal Wedding crowds making their way along The Mall. © iStockphoto.com/Matthew Dixon

	Event	*Demonstration*
Static	celebration, inauguration, commemoration	rally, sit-in
Mobile	parade	march

Size matters, particularly in demonstrations. A demonstration is in support of a movement or a cause. The movement has a number of sympathisers. The perceived number of sympathisers is important. If it is seen to be large, then it becomes easier to recruit others to the cause, and to persuade inactive sympathisers to join the demonstrations. Also, if the movement numbers are large then it is more difficult for authorities to deny their demands. It is as if the merit of the cause is somehow related to the number of its sympathisers.

People often judge an event's success by the size of the crowd. After a large demonstration, the crowd size estimates vary considerably. One side boosts the number to justify the cause, the other diminishes the number to weaken it. Journalists are often left somewhere in the middle, though the proximity of a newspaper's estimates to either extreme may give a measure of its political leaning.

Even in crowd estimation at public events, there may be an exaggeration factor at work in order to increase the status of the event, or of the individuals whom the event is honouring.

So estimating crowd size at public events has become much more about public relations and point-scoring than about a quest for the truth. But, even when we are after the truth, a crowd estimate tends to have a large error associated with it. As Steve Doig[1] said, writing about the numbers who gathered in Washington for President Obama's inauguration, "when it comes to accurately counting crowds, the slogan should be 'No, we can't.'" In the absence of accurate estimates, guesswork and point-scoring thrive.

Political point-scoring can be serious

The Million Man March was a gathering of social activists in Washington in October 1995. Because of the name of the event, the crowd count was crucial as a measure of its success. This led to extreme reactions to the resulting crowd estimates. March organisers estimated the crowd size at between 1 500 000 and 2 000 000 people. The United States Park Police estimated the crowd size at 400 000 and, as a result, the National Park Service was threatened with legal action.

Some heat was taken out of the debate by an intermediate estimate given by the Remote Sensing Unit at Boston University: 870 000 people, with a margin of error of about 25 per cent. Nevertheless, the Park Service estimate was never retracted, and the estimate has been supported by other statisticians. Following the controversy surrounding the Million Man March, the Park Police no longer give official crowd size estimates.

Estimates obtained by police or authorities are primarily aimed at crowd management and the provision of facilities and support. Such estimates should not be politically biased. But this is often not the public perception. It was once the case that newspaper estimates were regarded as unbiased. Many "historical" crowd figures are based on journalistic estimates. These days there is rather less enthusiasm for newspaper estimates.

Even when there is no political agenda crowd estimates can be quite variable. For example, estimates of the size of the crowd at the recent Royal wedding of Prince William and Kate Middleton ranged from 500 000 to 1 000 000. At the Obama inauguration ceremony, unofficial government estimates put the crowd at 1.8 million, based on information collected from satellite images, photos from a balloon and from individuals on the ground. Others[2], using the same photographic images, gave estimates closer to 1 million.

In the future, satellites will generate high-resolution satellite photos that will reveal enough detail for accurate crowd counting to be possible; they will be based on scanning programs that analyse the digital photo information. Judging by the Obama inauguration results, we are not there yet! Also, it may be some time before this information is generally available; the satellite images used for the Obama inauguration crowd came from a military satellite. There is still something to be gained from good on-the-ground data.

Static crowd estimation

So how does one get good – or even adequate – on-the-ground data? Counting all the heads directly at a demonstration or rally is usually impossible. We need some other way. The basis of estimation of a static crowd is simple enough, in theory: it boils down to area multiplied by density. The method was originally put forward by Jacobs in 1967, to estimate the size of the Berkeley riots. The plaza where students gathered was marked into grid squares, so a simple way to estimate the crowd was to count the number of squares and estimate how many students were in each square on average. His density rule, which is still used today, is that in a loose crowd the density is about 1 person per square metre, a solid crowd has about 2 persons per square metre and very dense crowds have about 4 per square metre.

A very quick estimate can be obtained by observation. A better estimate could be obtained by counting a sample of grid squares in real time or, better, from photos. Such sampling would lead to a standard error as well as an estimate.

If the crowd area is regular, clearly delineated and visible to the camera, then multiplying by the average density will produce a crowd estimate. But often the area occupied by the crowd will not be regular, and parts of it will remain invisible to the camera: trees, buildings or darkness may hide parts of the crowd. These are like non-respondents in a survey. This adds to the uncertainty of the estimate.

Photography is the best basis for estimation of density and for estimation of area. There are, though, a vast number of ways in which this can be done. Depending on the quality of the images pixel-based approaches can work; or you can use feature-based approaches such as texture, edge-points, shape or head-counting.

From a good photographic image, we can obtain estimates and standard error for the crowd area A, and for the crowd density D (see box below). Results suggest that with current technology the relative standard error (RSE), $\text{se}(\hat{N})/\hat{N}$, is of the order of 10%, even with a moderate photographic image. Compatible estimates which are different by a factor of 2, as at the Royal wedding, would require an RSE of about 20%; this would be at the upper end of the error spectrum.

Based on photographs, defensible bounds for area can be obtained. Then, given plausible bounds for the average density, bounds for the sample size can also be found.

All this can be useful in deciding which among conflicting estimates is nearer the truth. In the recent candlelight vigil in Victoria Park in Hong Kong to mark the 22nd anniversary of the June 4th crackdown on the pro-democracy movement in Beijing, Reuters, BBC and other major news agencies simply used the organiser's estimate of 150 000. The police estimate was 77 000. In this case, the area is reasonably precisely known: A = 42 000 m^2. We estimated the average density to be slightly less than 2 persons per square

Relative standard error for crowd density

Assuming that area estimate $\hat{A}$ and crowd density estimate $\hat{D}$ are independent, at least approximately, we obtain, using the delta rule, an expression for the relative standard error:

$$\frac{\text{se}(\hat{N})}{\hat{N}} = \sqrt{\frac{\text{se}^2(\hat{A})}{\hat{A}^2} + \frac{\text{se}^2(\hat{D})}{\hat{D}^2}}.$$

Here $\hat{N}$ denotes the estimate of the number in the crowd N, and se denotes standard error. The delta rule is an approximation valid when the relative standard errors are small.

With a suitable crowd area and photograph, the estimate could be improved by stratification into high-density, medium-density and low-density regions, for example.

metre, corresponding to a solid crowd, and giving an estimate close to the police estimate. An estimate of 150 000 would require that the entire area was covered with people at mosh pit density close to 4 persons per square metre. This was clearly not the case.

A common "event" crowd is a sporting event or a concert. Crowds at such events are simple to count because entry to the event is controlled. People are counted as they pass through turnstiles. Larger public events, such as Royal weddings and New Year's Eve gatherings, are not so controlled: there are many "points of entry", none of them with turnstiles. If the number of entry points is not too large, then counting at each entry point can provide or help to improve a crowd estimate. Such counting is not so simple, however, even if it is supplemented by photos or videos. At this stage the crowd is mobile rather than static.

The anti-Article 23 protest in Hong Kong on 1st July 2003. © iStockphoto.com/Chi Chung Leung

Mobile crowd estimation

While the area × density method applies to counting an assembled crowd, it is not well suited to a mobile demonstration such as a march. A mobile demonstration tends to be more variable. Even if good photographs are available the area is generally more difficult to specify, and while the crowd density is less it tends to be much more variable. The crowd itself is more variable in time and in space: a march may have a disciplined, well-formed head and a ragged, spread-out tail, and then gather more compactly at the destination to hear the speeches. The crowd estimates are considerably more uncertain for a mobile crowd than for a static one. For example, in the "Stop the War" demonstration in London in February 2007, organisers estimated 60 000, whereas the Police gave a figure of 10 000: a six-fold difference in claim[3].

Mobile demonstrations can be quite large. Since the sovereignty change in Hong Kong in 1997 there have been demonstrations and/or celebrations on 1 July each year to mark it. In 2003, the change of the registration of Article 23, for dealing with subversives, had drawn much opposition from the Hong Kong community. Large demonstrations contributed to a reversal of the change: on 1 July 2003, the crowd estimates ranged from 350 000 (as quoted by the police) to 700 000 (as quoted by protesters). The generally accepted figure is 500 000, a little less than one-tenth of the population at the time. This event has continued annually. The size of the crowd continues to be used as a bargaining tool in negotiations with the local and mainland governments. There have been many controversies concerning the actual size of the demonstration, even though in recent years the crowd has not approached the size of the 2003 march.

Two "on-the-ground" estimating techniques have been applied to the 1 July demonstration. The task in this and similar demonstrations is challenging because we have to estimate the size of a mobile population in which people may join or leave the march at various points.

Inspection points may be placed at various points along the route. These are points at which the number of participants is "counted". A "count" really means an estimate of the number of people passing the observation point – that is, the number counted by observers at sampled time intervals during the demonstration. There is often some focal point of the demonstration, usually towards the end of the route. However, it is generally more difficult to "count" there. Near the focus, there is often a larger area such as a square which allows the crowd to spread somewhat, and its movement is slower and more turbulent. Consequently, one or more other inspection points are used.

The number of participants, N, in the demonstration is defined to be the number of people who entered the demonstration route during the specified "demonstration" time. We now describe two on-the-spot counting methods that have been used. Each method can give a standard error as well as an estimate for the crowd number N.

Count and follow-up

The method is to choose an inspection point P, generally near to the focus, and to "count" the number of participants passing this point. This gives us an initial estimate $\hat{N}_P$.

However, even if the counting method were accurate, there is still a problem: some participants may have left the route before our inspection point, or have joined the route after it, which would mean that our estimate is too low. One method used to overcome this is a subsequent random phone survey of the population. This involves first finding the participants, then asking them whether they passed our inspection point or not.

Even a big demonstration involves only a small proportion of the population, and so a large number of phone calls would need to be made. There are, of course, other problems related to a phone survey. These include possible non-representativeness of the phone sample, the truthfulness of the responses, an allowance for clustering and household groups, dealing with non-response and so on. And in a tense and politically charged situation, such as the demonstrations in Tunisia, Egypt and other countries that have led to the Arab Spring, those

Telephone and double-counting estimates of crowd size

Telephone follow-ups

If p denotes the proportion of participants who passed P in the march, we assume that the phone survey provides an estimate $\hat{p}$ with standard error given by

$$se(\hat{p}) = \sqrt{\hat{p}(1-\hat{p})/k},$$

where k denotes the number of participants responding to the phone survey. The count estimate is then given by $\hat{N} = \hat{N}_P / \hat{p}$, and using the delta method, we find the standard error is given by

$$\mathrm{se}(\hat{N}) = \sqrt{\frac{\mathrm{se}^2(\hat{N}_P)}{\hat{p}^2} + \frac{\hat{N}_P^2(1-\hat{p})}{k\hat{p}^3}}.$$

Double counting

If we have two observation points, A and B, and ignoring, or excluding, individuals who both entered and left the route between A and B, an estimate of N, the number of marchers, is given by

$$\hat{N} = \hat{N}_A + (1-\hat{q})\,\hat{N}_B,$$

where $\hat{N}_A$ and $\hat{N}_B$ denote the estimated counts at A and B, and $\hat{q}$ denotes an estimate of the proportion of individuals passing A who also passed B. The quantity $\hat{N}$ estimates the numbers of individuals who passed either A or B. A reasonable estimate of q can be obtained by asking a representative selection of m individuals passing B whether they joined the demonstration route before A. This yields an unbiased estimate of q with standard error $\sqrt{\hat{q}(1-\hat{q})/m}$. Again the standard error of $\hat{N}$ can be approximated using the delta method:

$$\mathrm{se}(\hat{N}) = \sqrt{\mathrm{se}^2(\hat{N}_A) + (1-\hat{q})^2\,\mathrm{se}^2(\hat{N}_B) + \hat{N}_B^2\,\frac{\hat{q}(1-\hat{q})}{m}}.$$

who were present may have well-justified fears about admitting it, on the phone, to a stranger.

Nevertheless, if a phone check is possible it can again give both an estimate and a standard error. The box above sets out the mathematics.

Double counting and spot-checking

Another method, which avoids the use of a phone survey, is to choose not one but two inspection points, *A* and *B*, on the route of the march, not too close together and with one close to the end.

Again there is a problem to be overcome: there will be participants who passed one but not both of the inspection points, and some who passed neither. This method increases the counting cost and adds the cost of the spot-check survey described below, but avoids the time and cost of the phone survey. It also avoids the phone response bias. This method has the advantage of increased efficiency and immediacy. The details are given in the box above.

A comparison of the standard errors using the two methods indicates that the two-inspection-point method gives a more accurate estimate. It is also more immediate: a same-day estimate is obtained.

In an application to the Hong Kong march of 2006, the two-inspection-point estimate was found to have a relative standard error of 2%[4]. Again, we found that our estimate was closer to and compatible with the official police estimate. The organisers' estimates have been consistently and substantially larger than the police estimate throughout the years, by a factor of 2 to 3. This year, in the march of July 1, 2011, the estimates were 60 000 to 220 000. It appears that the organisers' exaggeration factor is increasing.

Having increased the number of inspection points from one to two, and having found that some benefit ensued, one might ask if there might be further benefit in using more inspection points. However, more inspection points would involve a further increase in counting staff and more questionnaires, which would need to be more complicated and therefore more likely to cause respondent error. The result would be increased expense for little increase in precision. Rather than have more counting points, the issue is to choose "good" locations for the counting points – suitable points on the route where counting is feasible and can be done accurately.

Conclusion

In the end, the public would be better served by an accurate crowd estimate rather than an "advertising" estimate. When a crowd estimate is given based on the area × density method, values for the area and the density should be given. This should restrict the range of estimates and give a better basis for their comparison.

Crowd estimation is quite variable, even when it is based on photographic images. Image analysis is continually improving. It will provide accurate answers in the future, but it is less than perfect. There is still merit in on-the-ground data, particularly for mobile crowds.

Crowd estimation methodology is such that estimates with relative standard error of less than 10% can be achieved given reasonable technology. However, it seems there is too much politics in the mix for crowd estimation to be made precise in the near future. The public has a view of the truth that is coloured by their beliefs. This applies particularly to crowd estimation. If a demonstration is perceived to be in a good cause, then exaggeration of crowd size is regarded as a white lie, while underestimation is seen as politically motivated. Crowd estimation is viewed with scepticism. If it is not, then it should be.

References

1. Doig, S. (2009) How big will the inaugural crowd be? Do the math. http://www.msnbc.msn.com/id/28662672/bs/
2. Nevarez, G. (2009) Professor estimates crowds with satellite image. http://www.statepress.com/archive/node/3654
3. Rosenbaum, M. (2007) How many there? BBC News. http://www.bbc.co.uk/blogs/opensecrets/2007/08/how_many_there.html
4. Yip, S. F. P., Watson, R. *et al.* (2010) Estimation of the number of people in a demonstration. *Australian & New Zealand Journal of Statistics*, **52**, 17–26.

Ray Watson is Associate Professor in the Department of Mathematics and Statistics at Melbourne University. Paul Yip is Professor in the Department of Social Work and Social Administration at the University of Hong Kong.