Crowd size estimate based on line density integration and computerized count of photos

Albert Satorra (UPF) & Josep M Oller (UB)

Talk at the Multivariate Analysis Group of SEIO

Barcelona 20-21, October, 2016

Acknowledgement

This work was originated with the crowd counting by *Societat Civil Catalana* of the September 11 demonstrations in Barcelona of 2015 and 2016. For the crowd counting reported here we very are grateful to Juan Amor, Montserrat Baras, Eduardo Chibás, Felipe Moreno, José Rosiñol, Josep Vaquero, among others, for their altruistic work.

Special thanks are due also to Professor Mubarak Shah and Dr. Haroon Idrees, of the Center for Research in Computer Vision (CRCV), University of Central Florida, for their kidness in applying their algorithm *Counting in Extremely Dense Crowd Images* —altruistically and in a record time— to the images we sent to them on the September 11 of 2015 and 2016.

The crowd counting (CC) problem

- *CC is a technique used to count or estimate the number of people in a crowd, the size of the crowd (SC). For political rallies SC carries political significance so the actual SC reported is controversial.*

  "There is a large amount of variability in crowd estimates [SC] for two reasons: it is difficult to do, and there are strong motivations for getting it wrong!" (Watson and Yip, 2011, p. 104). [1]

- For example, in the demonstration of 11 Sept 2015 in Barcelona, the organizers claimed 2 million people while other agencies gave a SC of just above 0.5 million.
- Huge disparities among SCs evidence the presence of non-statistical bias in some —or maybe all[2] — of the estimates reported.

[1] Watson, R and P. Yip (2011). How many were there when it mattered? Estimating the sizes of crowds, *Significance*, pp. 104-107.

[2] Daniel Verdú, El Pais (12 Sept. 2016), *Todas las cifras son mentira*.

Area method (Jacobs, 1967)

The simple formulae:

$$SC = A \times d$$

where area A is partitioned in a grid of squares and d is a mean density (people/m^2) across squares.

Avenue crowd (AC)

When the crowd concentrates on an avenue (as the ones of the 11S in Barcelona, 2015, 2016). The area method relies on three basic parameters

- Length: L
- Width: W
- Density (persons/m^2): D

In 11 sept. 2015, undubitative L= 5.2 Kms, but W and D debated numbers. [regarding W, from wall to wall a bit less than 50 m. a width 35m seems reasonable; for D , 1.5 up to 4, with 4 would be too compressed crowd . . .

In Sept. 2016, undubitative L= 1.6 Kms, but W and D debated numbers . . .

Our proposal of SC estimate, when the crowd is AC

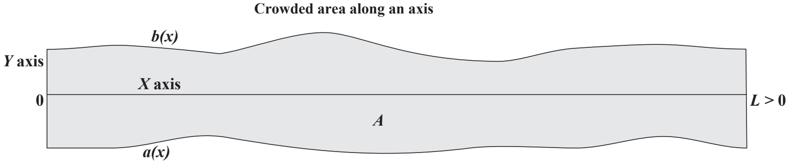
- It departures from the area method approach:
  it requires L but not W and D.
- Replaces the controversial parameters W and D by counts of people in aerial images taken randomly along the avenue.
- Automated Mass-Crowd Count obtains the number of people in each image
- *Google* /cartographic tools to compute the length of each image
- SC will be the integral of a line density estimated from the counts

True crowd size, *TCS*

- The (unknown) true parameter can be expressed as:

$$TCS := \iint_A \rho(x,y)dydx \tag{1}$$

where ρ is a continuous approx. of the density (people/m^2)



- Let $b(x)$ and $a(x)$ be limiting region A curves.
- If we define the *line density* $\rho_\ell(x)$ as:

$$\rho_\ell(x) \equiv \int_{a(x)}^{b(x)} \rho(x,y)\, dy \tag{2}$$

Fubini's Theorem

- Then by Fubini's Theorem:

$$TCS = \int_0^L \rho_\ell(x)\, dx \qquad (3)$$

- Photos are required, along the avenue, covering all the width of the occupied area in the avenue

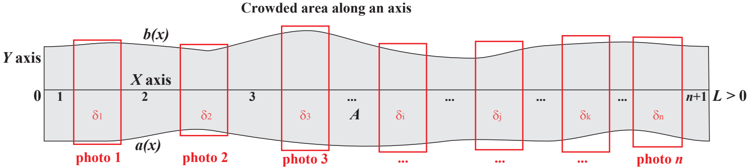


Photo sample

- Let $\delta > 0$ be a real positive number. Consider a photo which covers $[x, x + \delta] \times [a(x), b(x)]$. Let X be a uniform random variable on $[0, L]$. Then, the variable:

$$W_\delta(X) := \int_X^{X+\delta} \rho_\ell(t)\, dt \tag{4}$$

  shall be equal to the people registered on this photo, which depends on x.

- If $\delta << L$, as usual, the mean value of $W(X)$ will be given by:

$$E(W_\delta) = \int_0^L \left(\int_x^{x+\delta} \rho_\ell(t)\, dt \right) \frac{1}{L}\, dx \approx \frac{\delta}{L} \int_0^L \rho_\ell(t)\, dt = \frac{\delta}{L}\, TCS \tag{5}$$

TCS estimator

- Consider n randomly chosen photos along the axis $[0, L]$ of size n: $W_{\delta_1}(X_1), \ldots, W_{\delta_n}(X_n)$, and let us define a weighted average:

$$\widetilde{W}_n = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\delta_i} W_{\delta_i}(X_i) \tag{6}$$

then:

$$E(\widetilde{W}_n) = \frac{1}{L}\, TCS \tag{7}$$

- Therefore an estimator of *TCS* will be:

$$\widehat{TCS} = L\, \widetilde{W}_n \tag{8}$$

Notice that

sample values of each W_i are difficult to obtain through human eyes and therefore we shall use computer vision algorithms developed by H. Idrees y M. Shah, CRCV, UFC (USA)

TCS estimator with computer vision algorithms

- This computer algorithm supplies the maximum value of people, M_i, and the minimum, m_i, for each photo.
- If $X = x_i$, we shall consider that this computational result shall supply an estimate for $w_i = W(x_i)$, given by

$$\widehat{w}_i = \frac{m_i + M_i}{2} \qquad i = 1, \dots, n \tag{9}$$

This estimate carries on an error which we assume its approx. uniformly distributed among $[(m_i - M_i)/2, (M_i - m_i)/2]$. Therefore a combined estimate for the *TCS* will be:

$$\widehat{TCS} = \frac{L}{n} \sum_{i=1}^{n} \frac{(m_i + M_i)}{2\ \delta_i} \tag{10}$$

The accuracy of $\widehat{TCS}$

- The distribution of $\widehat{\tau}$ is not known, although we can expect by a version of CLT (Central Limit Theorem) approx. normally distributed.
- In any case we can estimate this distribution and, in particular, its standard error, through bootstrap or jacknife methods.
- Forcing non-overlapping, we deviate from a simple random sample, though accuracy of the crowd size estimates increases. Convenient numerical methods could be implemented to estimate the se. In the limiting case of not overlapping and covering all the axis, we shall have:

$$\sigma(\widehat{TCS}) = \frac{L}{\sqrt{12}\, n}\sqrt{\sum_{i=1}^{n}\frac{(M_i - m_i)^2}{\delta_i^2}} \tag{11}$$

A simulated crowd

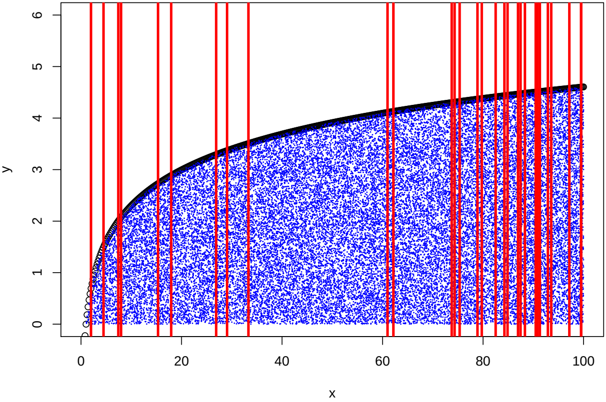


Figure: Simulated crowd and images (streps)

$\widehat{TCS}$ and simulation data

- We have the AC as the one depicted in Figure 1 The TSC is 360769
- We take 30 images along the main axis. We compute the exact number of points in each image; the density values are:

  `3076.333 2818.833 3041.833 4060.667 4379.667 4035.833 3445.833 3806.333 4578.500`

  `3163.833 4105.167 3328.167 4416.500 919.500 4405.333 4397.333 3149.000 4341.333 3818.167`

  `3195.833 3123.333 3047.333 2761.167 3752.500 4542.000 3652.167 4378.000 2714.667`

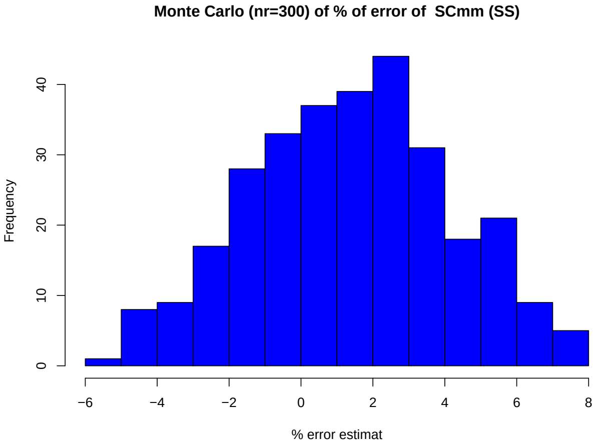
  `4299.667 4475.167`
- $\widehat{TCS}$ is the integral of the estimated line density.
  $\widehat{TCS} = 364100$.
  Absolute error is $364100 - 360769 = 3331$.
  Relative error $3331/360769$, below 1% .

Monte Carlo (replications 300)

- Relative error (in %) of the $\widehat{TCS}$

Min.	Q1	Median	Mean	Q3	Max.
-13.150	-4.132	-1.363	-1.271	1.672	10.090

sd = 4.41

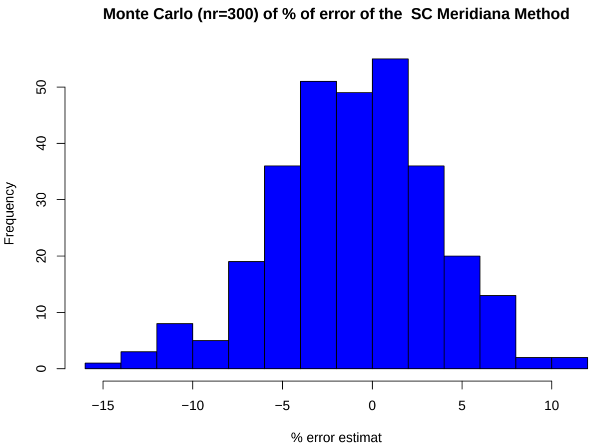


Monte Carlo evaluation (stratified sampling of images)

- Relative error (in %) of the $\widehat{TCS}$

Min.	Q1	Median	Mean	Q3	Max.
−5.0840	−0.6216	1.3050	1.3740	3.2920	7.8490

sd = 2.77



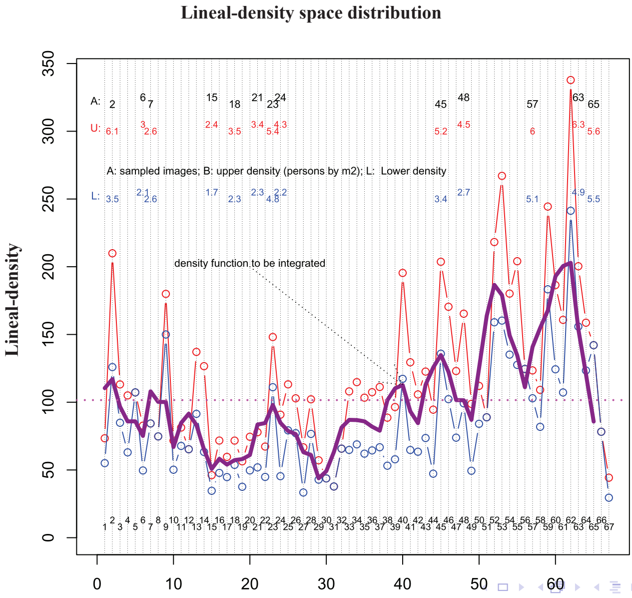
Via lliure 2015 (VL) count (SCC)

- Background previous work: size measurement of people on *Gigafoto*, *Via catalana 2013*.
- $n = 67$ photos covering about 4.564,24 m measured along the axis from a total of L=5.223 m. Counts (lower and upper bounds) for each photo provided by Haroon Idrees.
- If the average lineal-density estimates based on m_i, M_i, o its average $(m_i + M_i)/2$, are denoted as $\rho_{*\ell}$, ρ^*_ℓ and $\widetilde{\rho}_\ell$, we have the summary table:

Statistics	$\rho_{*\ell}$	ρ^*_ℓ	$\widetilde{\rho}_\ell$
Minimum	29,53	37,81	36,91
Maximum	241,30	337,81	289,56
Mean	84,44	119,82	102,13
Standard deviation	41,57	58,49	49,14
Range	211,77	300,00	252,65

Table: Descriptive statistics of $\rho_{\ell i} = W_i/\delta_i$.

Lineal-density VL estimation (SCC)



Final (SCC) estimate *Via lliure 2015*

$$TCS_* = 441045 \leq TCS \leq TCS* = 625842 \tag{12}$$

and *TSC* estimated:

$$\widehat{TCS} = 533443 \tag{13}$$

Rounding this number, we finally give:

$$\widehat{TCS} = 530000 \pm 95000 \;\; \text{people} \tag{14}$$

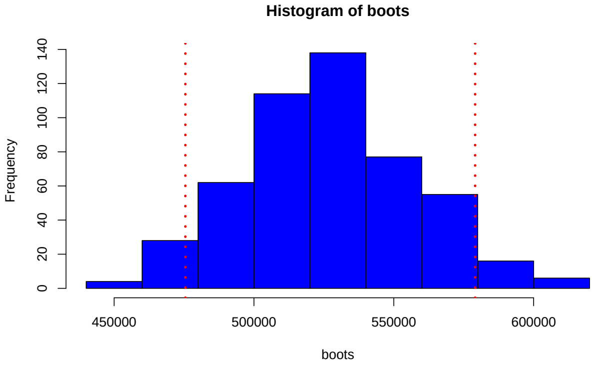


Figure: Bootstrap distribution (bs = 500) for the SC estimate of 11 Sept. 2015

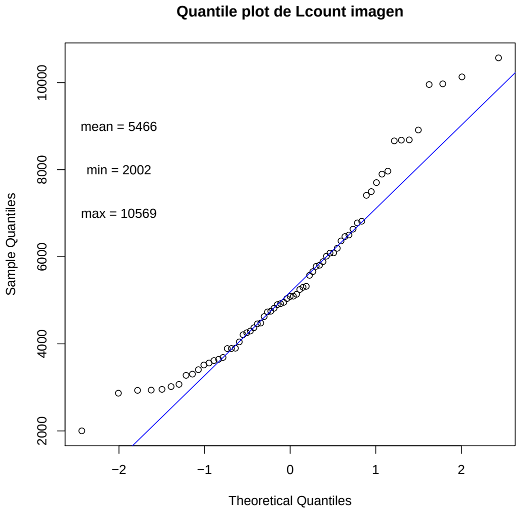


Figure: qqplot for the L count images

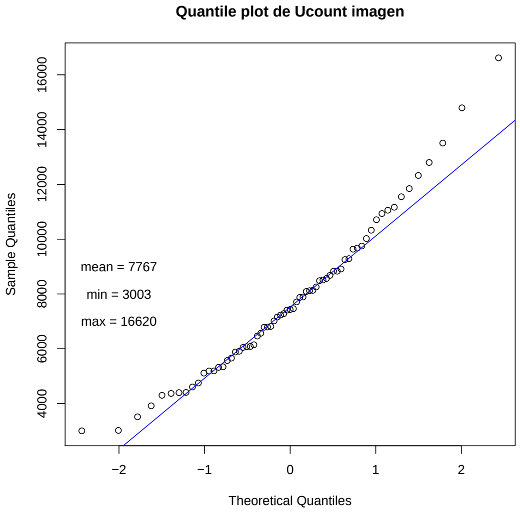


Figure: qqplot for the U count images

11 Sept. 2016, Passeig de St. Joan

- $n = 30$, L=1.8 Km, Images extracted from TV3 video of aerial view of the crowd (at a time close to the 5:14 pm pick time)
- If the average lineal-density estimates based on m_i, M_i, o its average $(m_i + M_i)/2$, are denoted as $\rho_{*\ell}$, ρ^*_ℓ and $\widetilde{\rho}_\ell$, we have the summary table:

Statistics	$\rho_{*\ell}$	ρ^*_ℓ	$\widetilde{\rho}_\ell$
Minimum	12.27	19.80	34.65
Maximum	59.88	87.11	145.20
Mean	26.62	39.78	66.41
Standard deviation	10.83	15.79	26.33
Range	47.61	67.31	110.55

Table: Descriptive statistics of $\rho_{\ell i} = W_i/\delta_i$.

$\widehat{TCS}$ 11 Sept. 2016, Passeig de St. Joan, 30 images

- Lower $\widehat{TCS}$: `95,840`
- Upper $\widehat{TCS}$: `143,219`
- Combining the L and U: `119,529`

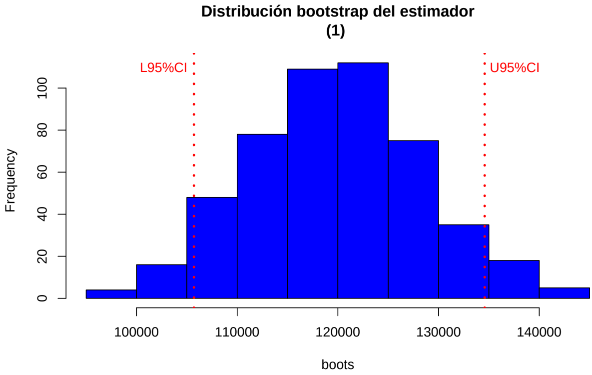


Figure: Bootstrap distribution for $\widehat{TCS}$, 11 Sept. 2016

Computation by area method of group OEC of SCC, crowd in Barcelona

- the rounded upper bound given by SCC based on the area method was: 140,000
- documentos Doc. 1
- documentos Doc. 2

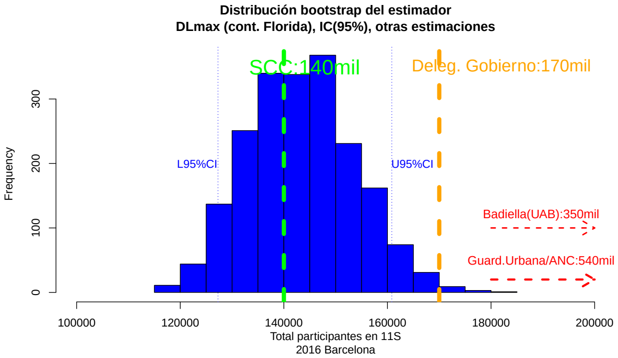


Figure: Comparative of SCs with the Bootstrap distribution for U based $\widehat{TCS}$ of 11 Sept. 2016, Barcelona crowd

Discussion

- A new estimate $\widehat{TCS}$ has been introduced for an static crowd in an avenue
- The procedure relies on Automated Mass-Crowd Count for images for a sample of images taken along the main axis of the crowd. The computer vision algorithms developed by H. Idrees y M. Shah, CRCV, UFC (USA) are essential for the new method.
- From the counts and length of each image, $\widehat{TCS}$ is the integral of a line density.
- $\widehat{TCS}$ overcomes (non-statistical) bias likely to arise using the classical area method.
- The new method is illustrated with simulated data and has been used by OEC of Societat Civil Catalana in crowd counting of the demonstrations in Barcelona of Sept 11, 2015 and Sept 11, 2016.

References

- Idrees H. Counting in Extremely Dense Crowd Images, Center for Research in Computer Vision, UCF, 2015.
- Watson, R and P. Yip (2011). How many were there when it mattered? Estimating the sizes of crowds, *Significance*, sept. issue, pp. 104-107. DOI: 10.1111/j.1740-9713.2011.00502.x
- Documents of Sociedad Civil Catalana:
  - 2013: Informe sobre el recuento de individuos presentes en la *Gigafoto* de la Via Catalana
  - 2015: Informe sobre el recuento de individuos presentes en la Via, SCC, 2015
  - 2016: Recuento de asistentes a las concentraciones de la Diada del 11 de Sept. de 2016
    1. Area method (all concentrations in Catalonia) Doc. 1, SCC, 2016
    2. line density method (crowd of Barcelona only) Doc. 2, SCC, 2016
- Jacobs, H. (1967). To count a crowd. *Columbia Journalism Review*. *6*, 36-40.

Haroon Idrees

Haroon Idrees and Barcelona