

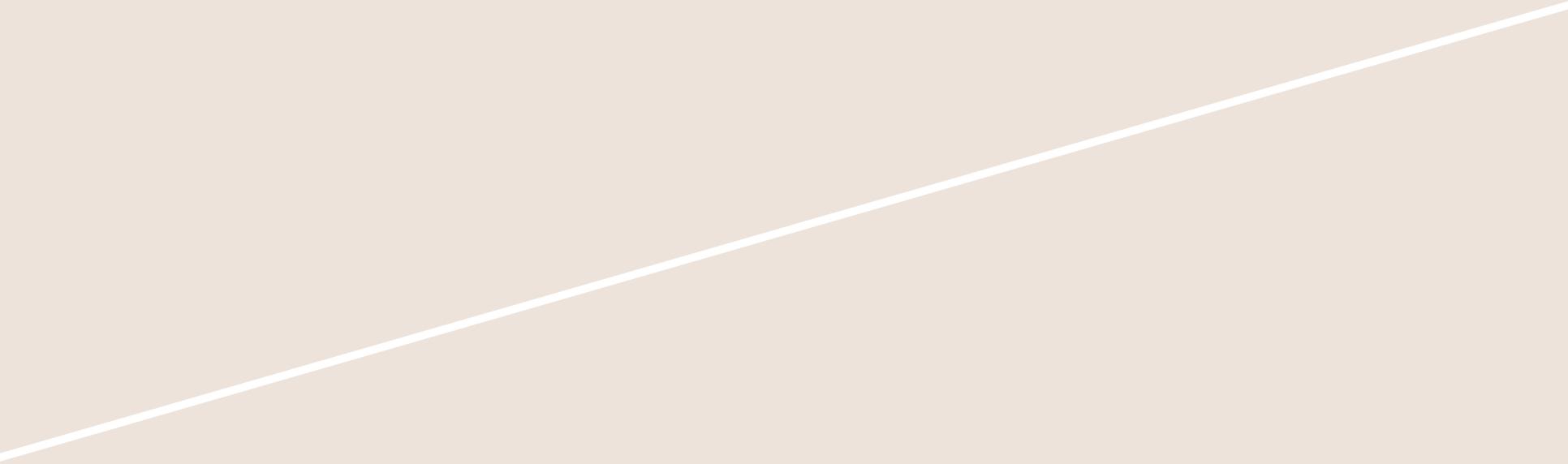
# APPLICAZIONE DELLA DISTRIBUZIONE DI POISSON:

Bombardamenti su Londra

MODELLI E METODI DELL'INFERENZA STATISTICA

Niccolò Foglia  
Michelangelo Garroni  
Federico Lancellotti  
Giacomo Lorenzon

# Introduzione storica



# Contesto storico

- Londra, 13 giugno 1944
- Una nuova arma “segreta”: la V1





## AN APPLICATION OF THE POISSON DISTRIBUTION

By R. D. CLARKE, F.I.A.

*of the Prudential Assurance Company, Ltd.*

READERS of Lidstone's *Notes on the Poisson frequency distribution* (F.I.A. Vol. LXXI, p. 284) may be interested in an application of this distribution which I recently had occasion to make in the course of a practical investigation.

During the flying-bomb attack on London, frequent assertions were made that the points of impact of the bombs tended to be grouped in clusters. It was accordingly decided to apply a statistical test to discover whether any support could be found for this allegation.

An area was selected comprising 144 square kilometres of south London over which the basic probability function of the distribution was very nearly constant, i.e. the theoretical mean density was not subject to material variation anywhere within the area examined. The selected area was divided into 576 squares of  $\frac{1}{4}$  square kilometre each, and a count was made of the numbers of squares containing 0, 1, 2, 3, ..., etc. flying bombs. Over the period considered the total number of bombs within the area involved was 537. The expected numbers of squares corresponding to the actual numbers yielded by the count were then calculated from the Poisson formula:

$$Ne^{-m}(1 + m + m^2/2! + m^3/3! + \dots),$$

where

$$N = 576 \quad \text{and} \quad m = 537/576.$$

The result provided a very neat example of conformity to the Poisson law and might afford material to future writers of statistical text-books.

The actual results were as follows:

No. of flying bombs per square	Expected no. of squares (Poisson)	Actual no. of squares
0	226.74	229
1	211.39	211
2	98.54	93
3	30.62	35
4	7.14	7
5 and over	1.57	1
	576.00	576

The occurrence of clustering would have been reflected in the above table by an excess number of squares containing either a high number of flying bombs or none at all, with a deficiency in the intermediate classes. The closeness of fit which in fact appears lends no support to the clustering hypothesis.

Applying the  $\chi^2$  test to the comparison of actual with expected figures, we obtain  $\chi^2 = 1.17$ . There are 4 degrees of freedom, and the probability of obtaining this or a higher value of  $\chi^2$  is .88.

# Contributo statistico di Clarke

- Clarke divise in riquadri un'area di 144 km<sup>2</sup> e contò il numero di colpi V1 in ogni singolo quadrato.

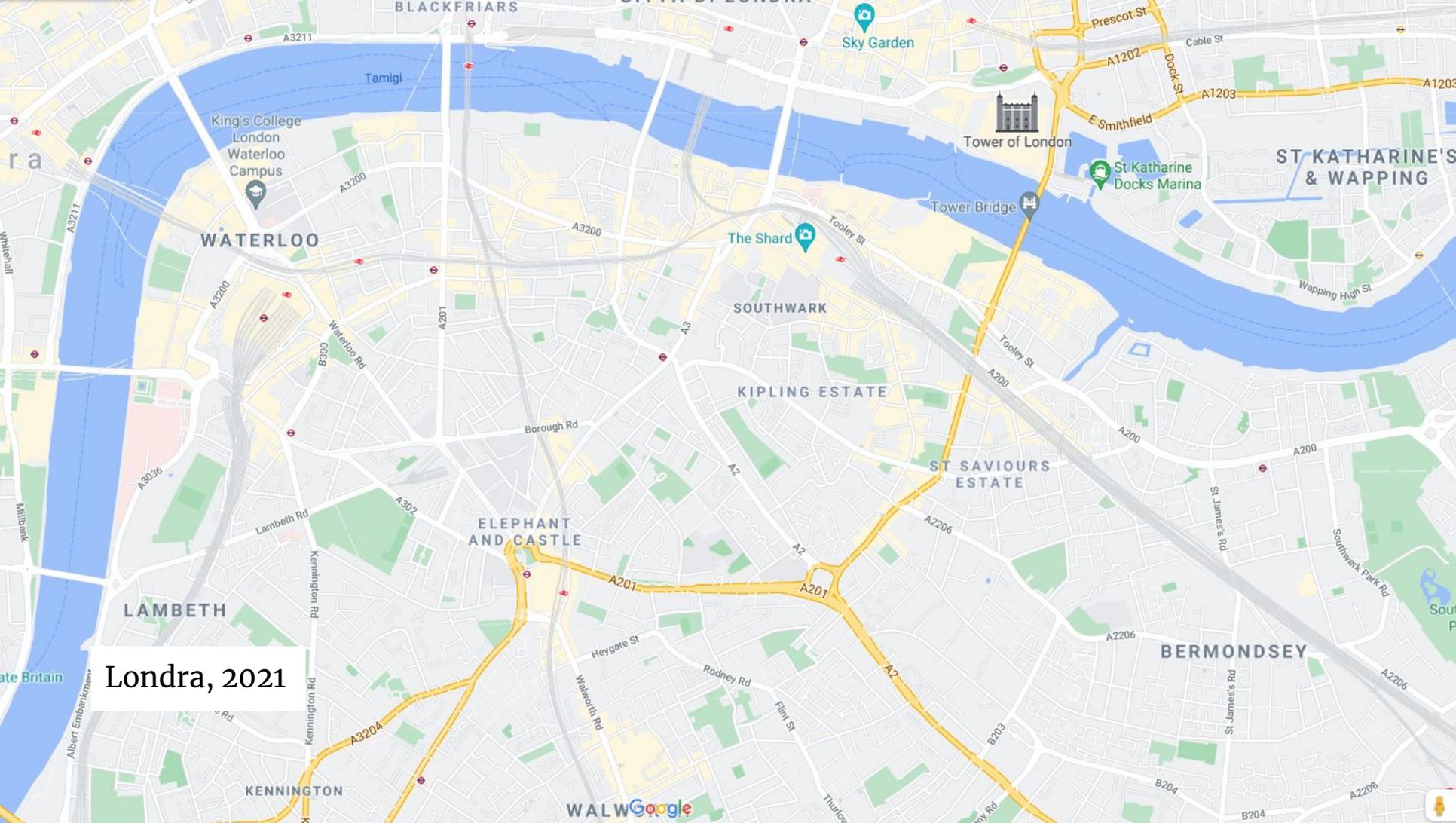
Numero V1 per riquadro	Numero atteso di riquadri	Numero effettivo di riquadri
0	226.74	229
1	211.39	211
2	98.54	93
3	30.62	35
4	7.14	7
5 o più	1.57	1
	576.00	576

# Rappresentazione dati e prima analisi

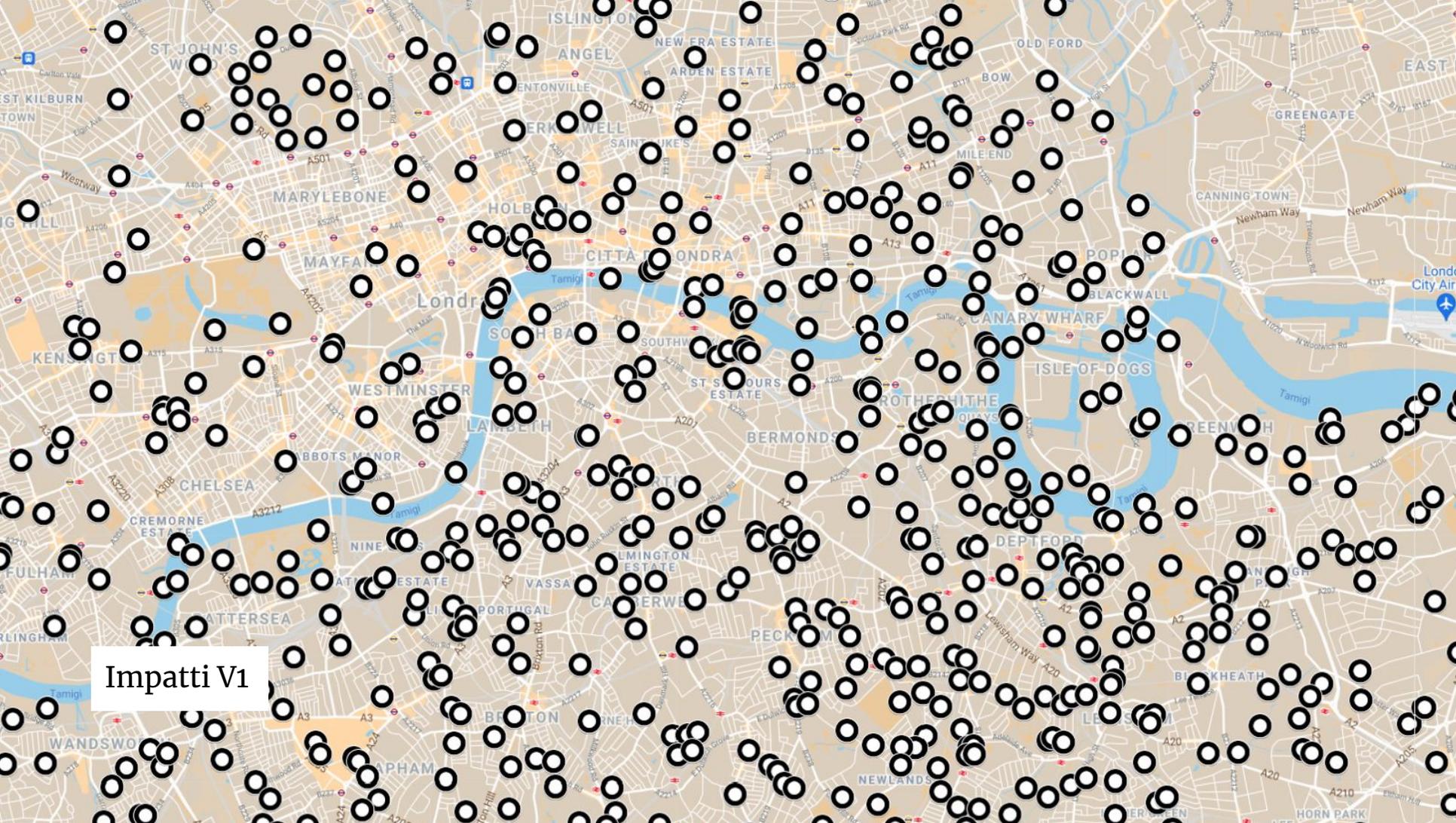




Londra, 1946



Londra, 2021

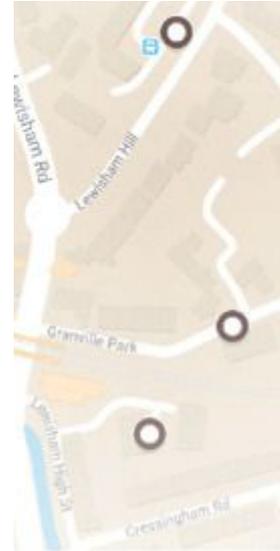


Impatti V1

# Cluster ?

Un gruppo di tre bombe V1 vicino a una stazione nel sud di Londra.

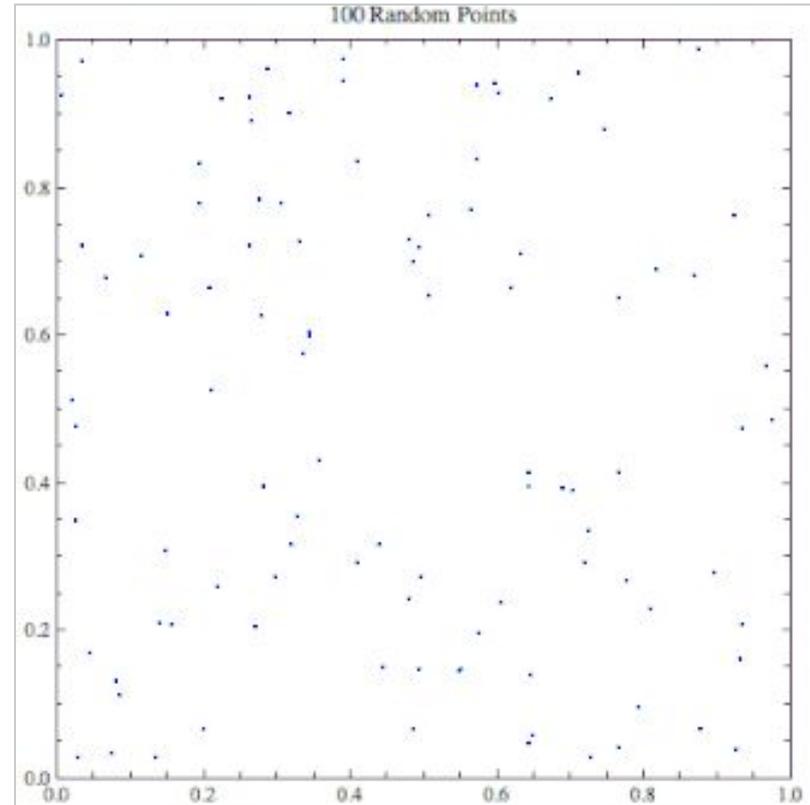
Cluster simili su posizioni strategiche fecero temere una grande precisione balistica tedesca.



# “Clustering illusion”

Tendenza a individuare cluster in distribuzioni spaziali uniformi

L'illusione è causata da una tendenza umana a sottovalutare la quantità di variabilità che potrebbe apparire in un piccolo campione di dati casuali.



# Obiettivo

Trovare evidenza statistica per escludere che i tedeschi possedessero un sistema di puntamento avanzato.

I cluster sono conseguenza di una distribuzione uniforme.

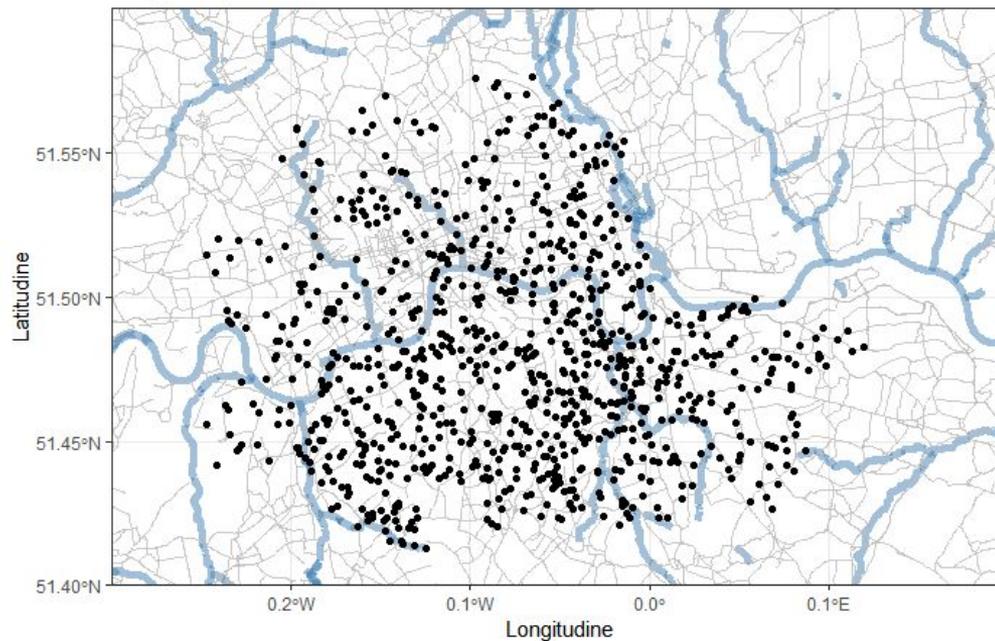
# Visualizzazione dati

```
LATITUDINE  LONGITUDINE
1  51.56940  -0.14715
2  51.57535  -0.09714
3  51.57390  -0.08524
4  51.57266  -0.08632
5  51.56940  -0.07931
6  51.56996  -0.06914
...
```

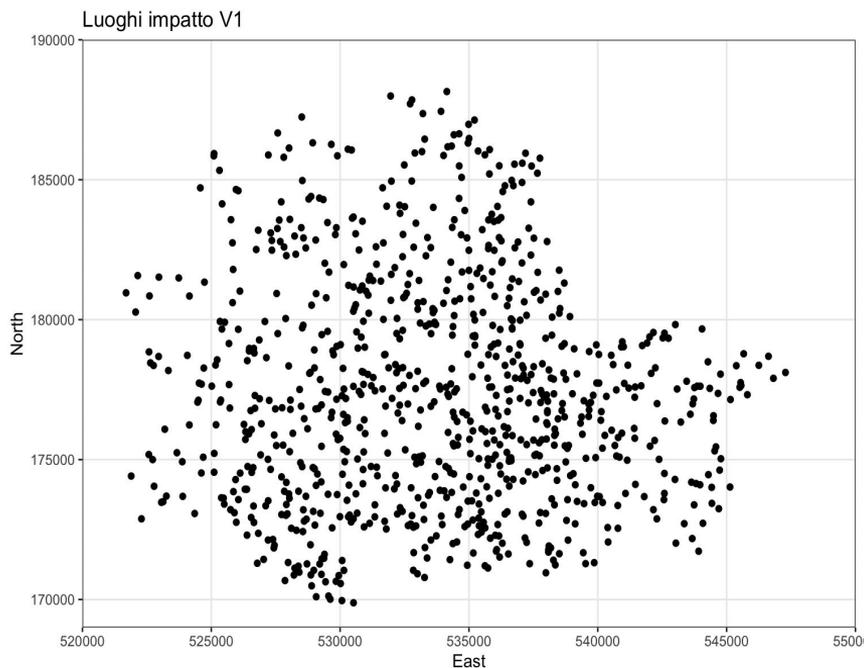
```
LATITUDINE  LONGITUDINE
Min.   :51.41  Min.   :-0.24765
Max.   :51.58  Max.   : 0.11973
```

```
dim( coordinate )
[1] 887  2
```

Luoghi impatto V1



# Conversione coordinate



Conversione in coordinate **UTM**  
(Proiezione Universale Traversa di Mercatore)

- Coppia numero-lettera (settore di riferimento)
- Due valori numerici (east e north)

**EAST:** quanti metri dobbiamo spostarci rispetto al meridiano centrale che passa per quel settore

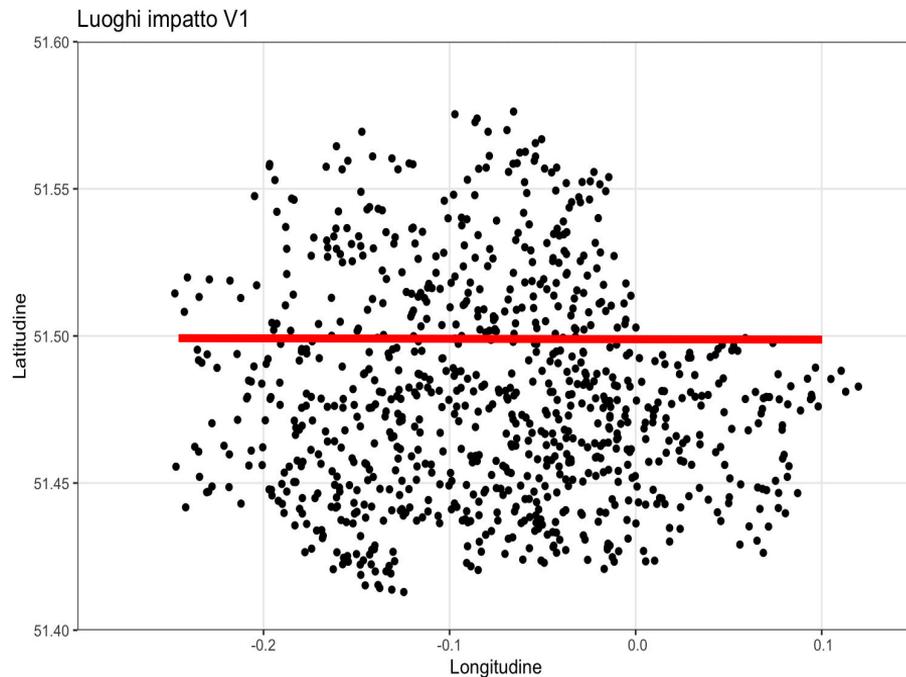
**NORTH:** quanti metri dobbiamo spostarci rispetto all'equatore.

# Differenza Nord-Sud

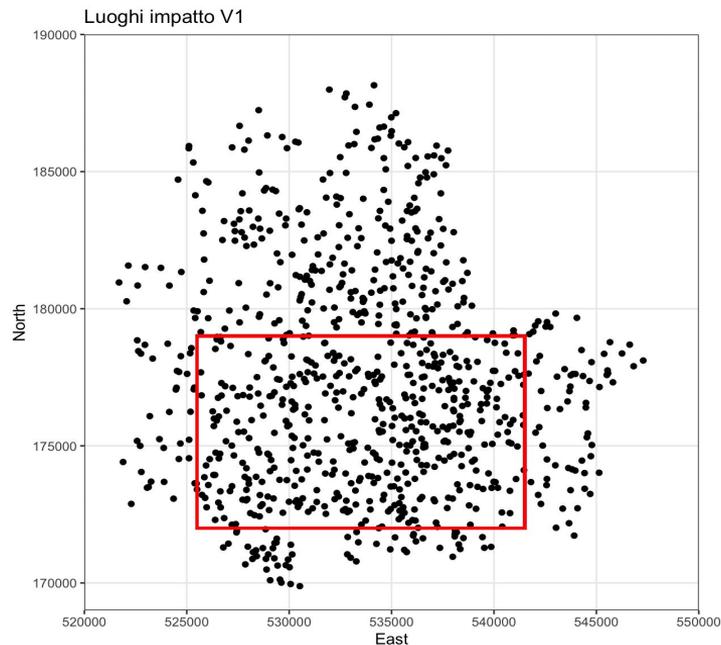
Selezione subset Londra-sud.

Motivi storici:

- “fall short” delle V1;
- bassa densità di popolazione;
- controspionaggio.



# Griglia



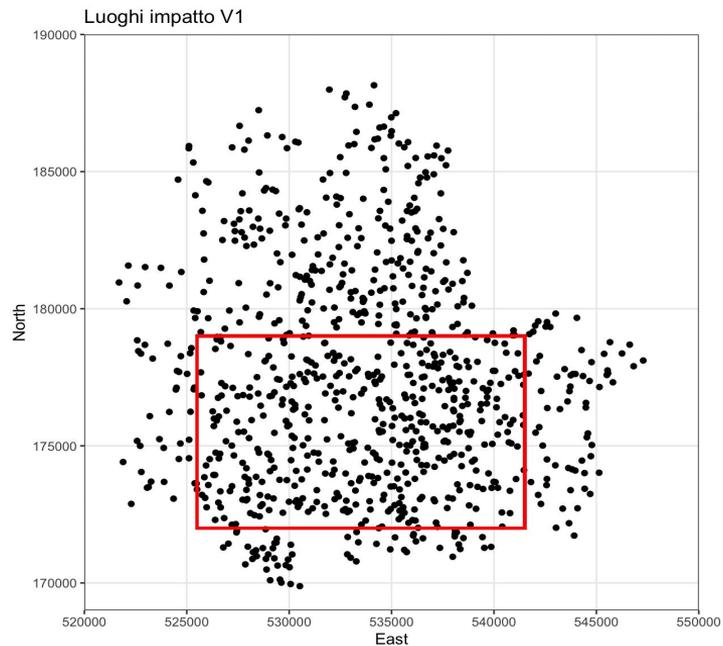
Area totale coperta: circa 112 km<sup>2</sup>

```
> dim(grid)
[1] 32 14
```

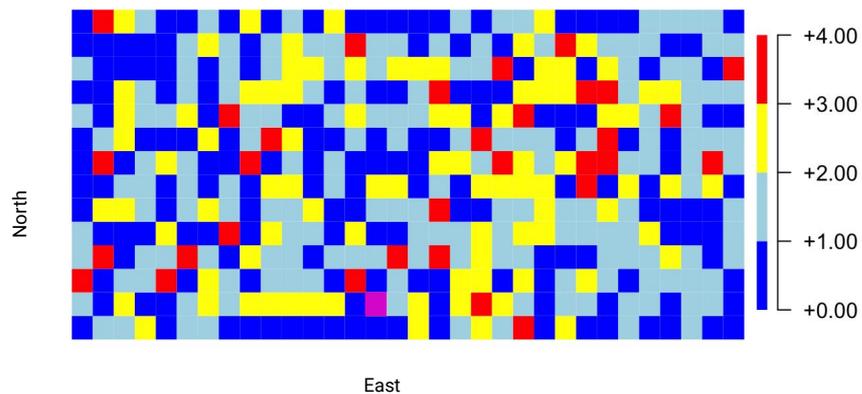
```
> dim(coordinate)
[1] 446 2
```

$$\lambda = 446/448 = 0.9955357$$

# Griglia



## Numero di impatti di V1 per settore



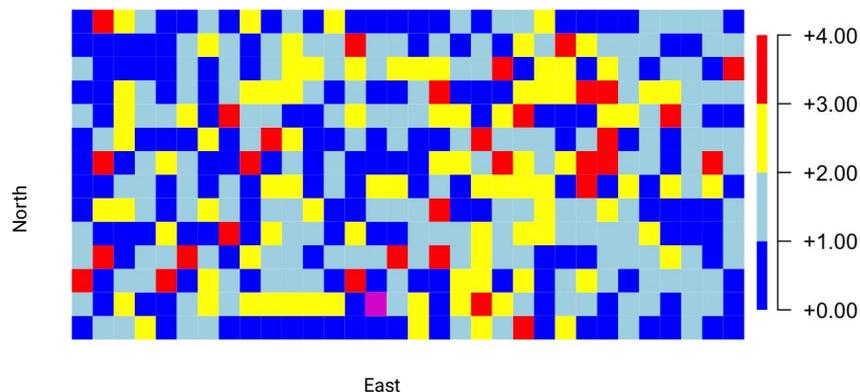
# Modello

$$P(n, \lambda) = \frac{\lambda^n}{n!} e^{-\lambda}$$

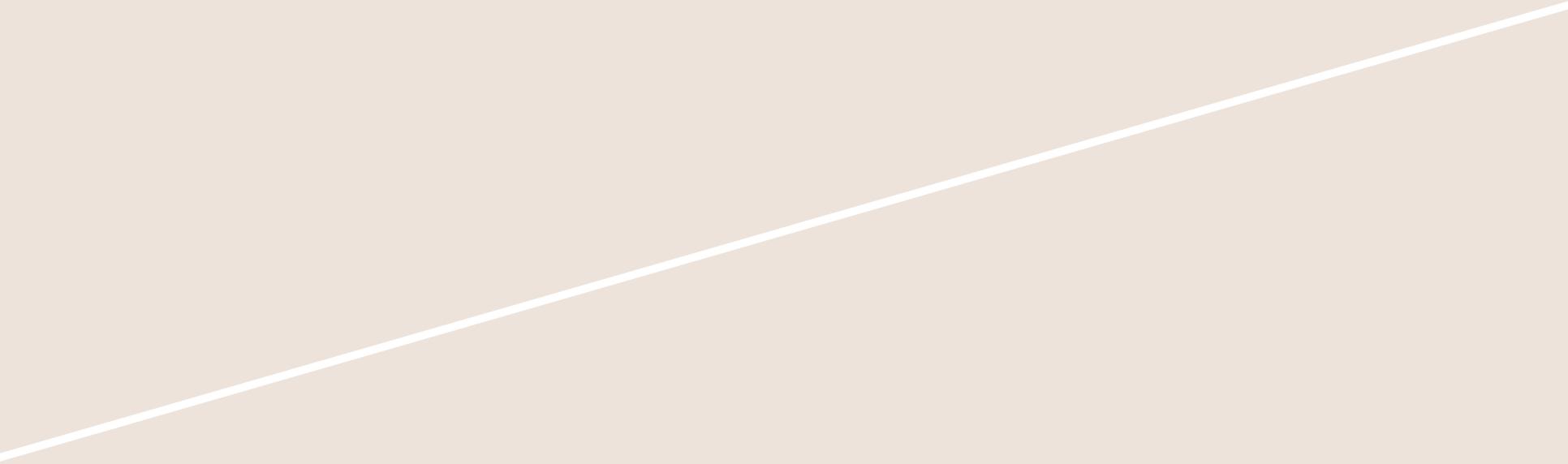
$$E(X) = \lambda \quad \text{Var}(X) = \lambda$$

Per  $n$  arbitrario di impatti, la distribuzione di Poisson modella il numero di settori in cui impattano quante bombe (0, 1, 2, 3, ...)

**Numero di impatti di V1 per settore**



# Analisi



# Test Chi-Quadro Buon Adattamento

$$\chi^2 = \sum_{i=1}^N \frac{(E_i - O_i)^2}{E_i}$$

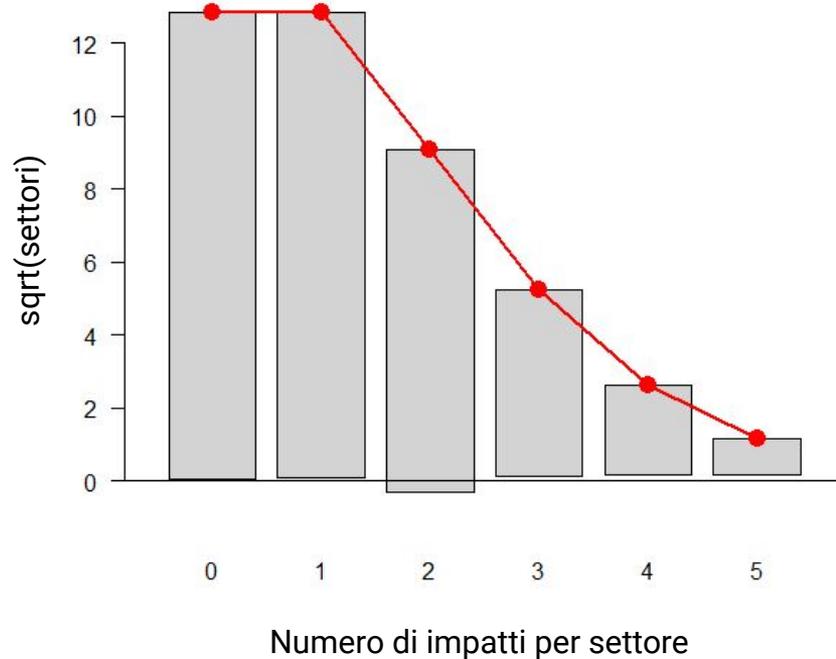
$E_i$  = valori teorici

$O_i$  = valori osservati

- Ipotesi nulla: il campione è estratto da una popolazione di cui è nota la distribuzione
- Test non parametrico
- N-1-m gradi di libertà, m uguale al # di parametri stimati per calcolare i valori teorici

# Procedura

- Individuazione area di analisi
- Costruzione griglia
- Conteggio del numero di missili per singolo settore
- Calcolo delle frequenze assolute e relative
- Stima dei parametri di Poisson
- Calcolo dei valori teorici
- Test Chi quadro



Primo test di buon adattamento		
#V1 per quadrato	Frequenza Osservata	Frequenza Teorica
0	164	165.547
1	163	164.808
2	88	82.036
3	26	27.223
4	6	6.775
5 o oltre	1	1.609
Totale	448	448

$$\lambda = 0.9955357$$

Chi-Squared: 0.8419731

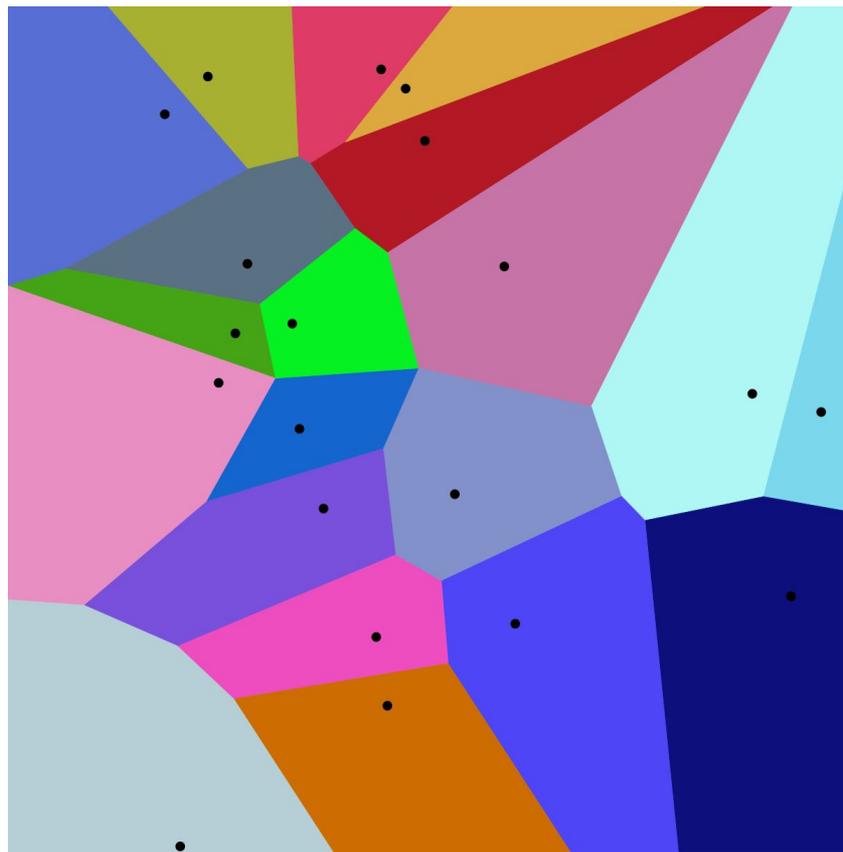
Gradi di libertà: 4

**P-value: 0.933**

Risultati del primo test di buon adattamento

# Procedura Griglia di Voronoi

- Costruzione dei diagrammi di Voronoi
- Conteggio del numero di impatti per tassello
- Calcolo delle frequenze assolute e relative
- Stima dei parametri della Poisson
- Calcolo dei valori teorici
- Test chi quadro
- Più iterazioni



Layer uniforme - VORONOI

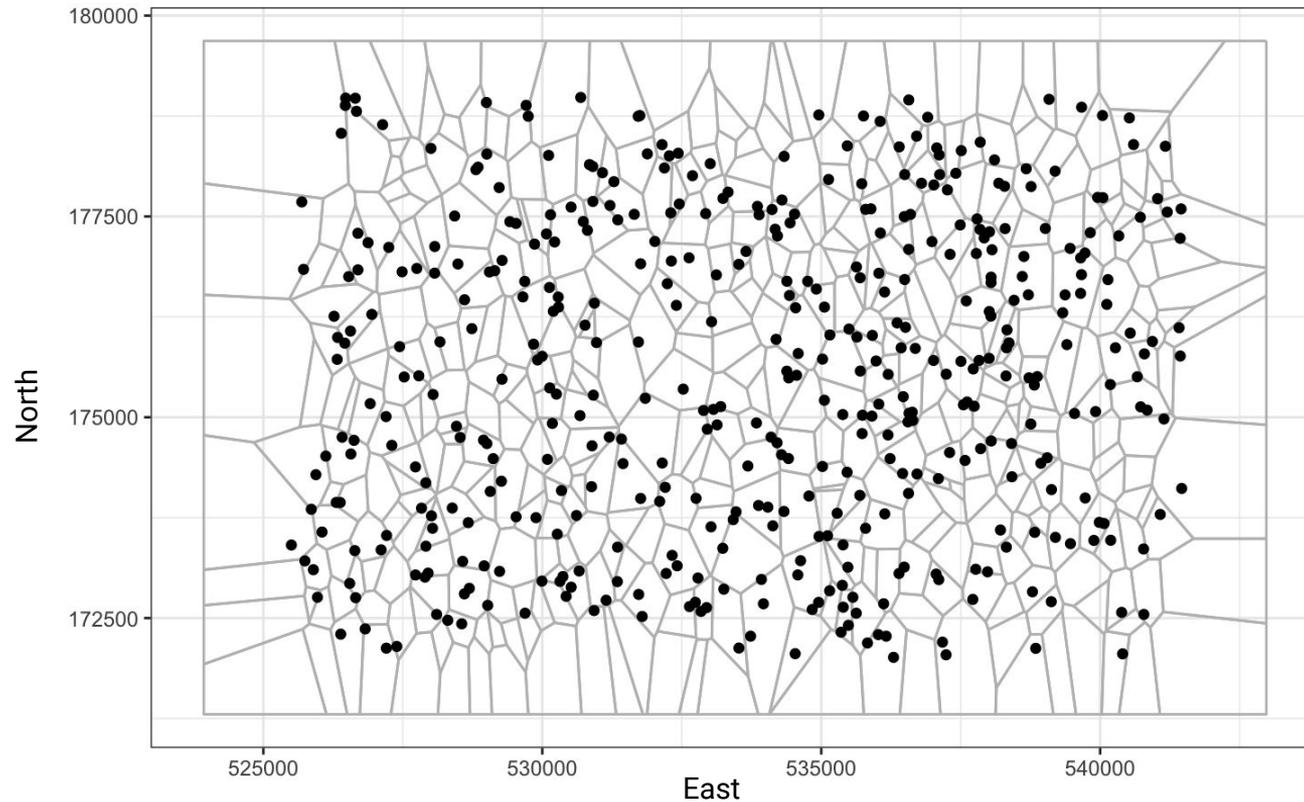


Diagramma di Voronoi - 448 punti casuali.

Test con griglie di <b>Voronoi</b>					
#V1 per tassello	Frequenza osservata (iterazione 1)	Frequenza osservata (iterazione 2)	Frequenza osservata (iterazione 3)	Frequenza osservata (iterazione 4)	Frequenza teorica
0	182	186	186	178	165.547
1	137	141	140	153	164.808
2	88	77	75	82	82.036
3	33	31	35	21	27.223
4	6	9	10	8	6.775
5	2	2	1	3	1.349
6 o oltre	0	2	1	3	0.260
Chi-squared	8.649539	21.97409	13.40248	5.447206	
<b>p-value</b>	<b>0.124</b>	<b>0.001</b>	<b>0.020</b>	<b>0.244</b>	

Risultati del test di buon adattamento utilizzando le griglie di Voronoi (448 tasselli)

Layer uniforme - VORONOI

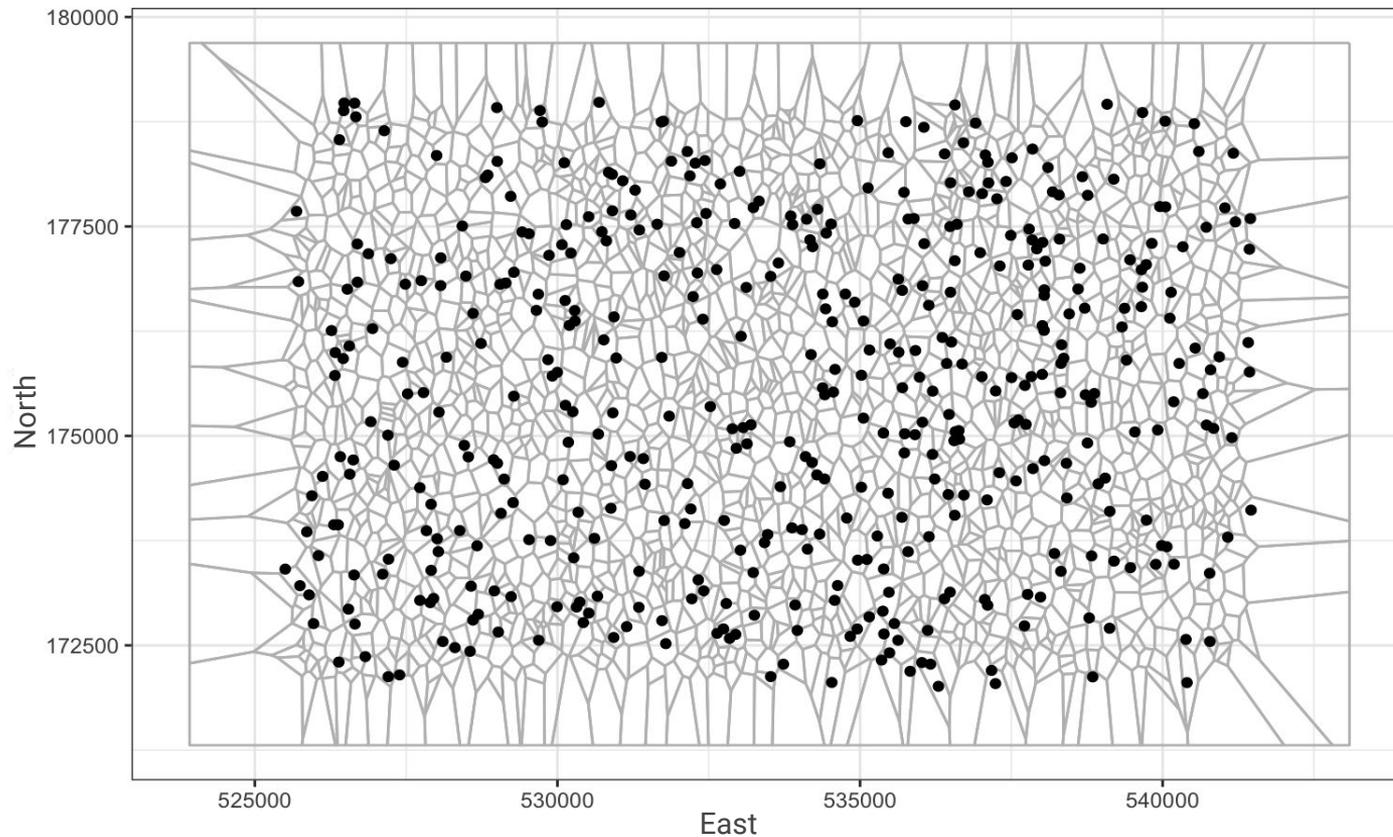
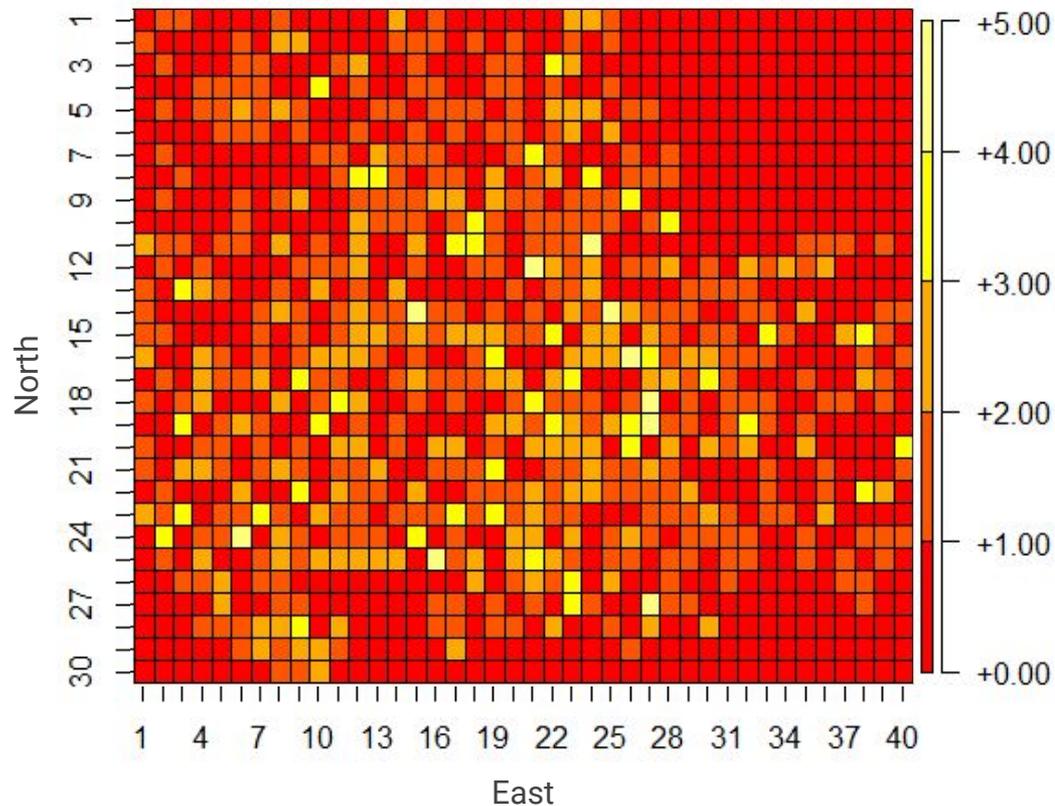


Diagramma di Voronoi - 2000 punti casuali.

Test con griglie di <b>Voronoi</b>					
#V1 per tassello	Frequenza Osservata (iterazione 1)	Frequenza Osservata (iterazione 2)	Frequenza Osservata (iterazione 3)	Frequenza Osservata (iterazione 4)	Frequenza Teorica
0	1609	1604	1607	1613	1600.229
1	339	351	347	336	356.851
2	49	40	41	44	39.788
3	3	5	3	6	2.957
4 o oltre	0	0	2	1	0.172
Chi-squared	3.074025	1.516265	0.3380661	4.895454	
<b>p-value</b>	<b>0.215</b>	<b>0.469</b>	<b>0.844</b>	<b>0.086</b>	

Risultati del test di buon adattamento utilizzando le griglie di Voronoi (2000 tasselli)

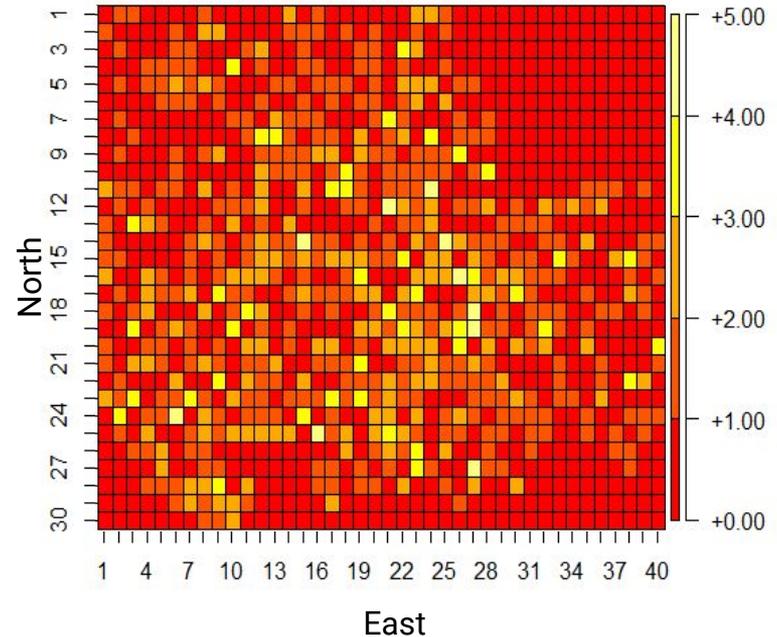


Numero di impatti per settore

# Distribuzione ZIP

(Zero Inflated Poisson)

- Versione modificata di una distribuzione di Poisson
- Permette di considerare opportunamente le ampie aree prive di osservazioni
- Possiamo applicare il modello all'intera mappa



# Test di buon adattamento con la distribuzione ZIP

- Analisi dei limiti del dataset
- Costruzione griglia **su tutta l'area**
- Conteggio degli impatti per singolo settore
- Calcolo delle frequenze assolute e relative
- Stima dei parametri della ZIP
- Calcolo dei valori teorici utilizzando la distribuzione ZIP
- Test di buon adattamento

# Distribuzione ZIP

(Zero Inflated Poisson)

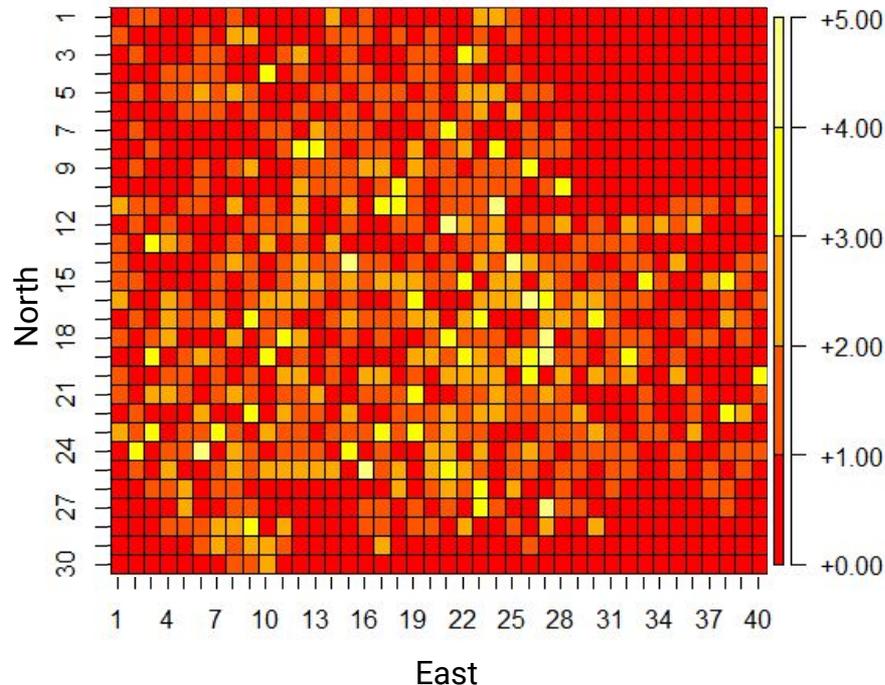
$$P(X = k) = \begin{cases} \pi + (1 - \pi) \exp(-\lambda) & \text{if } k = 0 \\ (1 - \pi) \exp(-\lambda) \lambda^k / k! & \text{if } k \in \{1, 2, \dots\} \end{cases}$$

$$\hat{\lambda}_{mo} = \frac{s^2 + m^2}{m} - 1$$

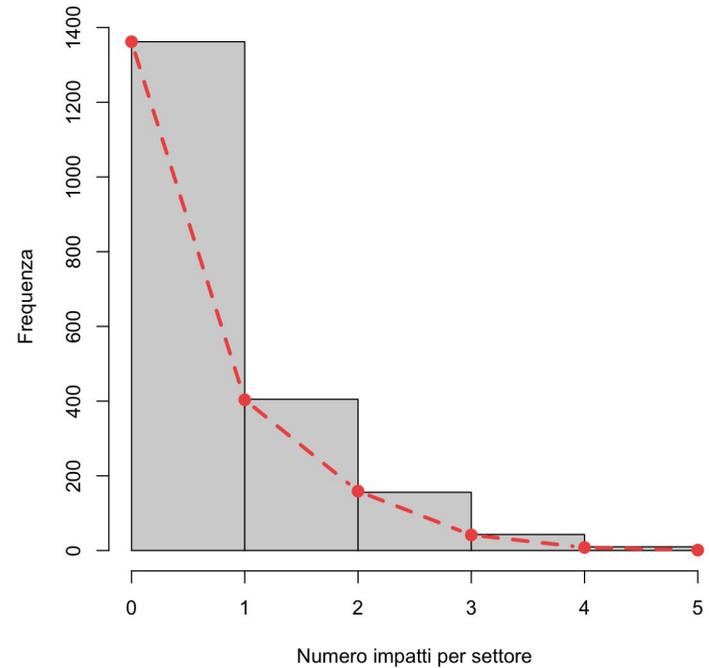
$$\hat{\pi}_{mo} = \frac{s^2 - m}{s^2 + m^2 - m}$$

$$E(X) = \lambda(1 - \pi)$$

$$\text{Var}(X) = \lambda(1 - \pi)(1 + \lambda\pi)$$



Test con distribuzione ZIP		
#V1 per quadrato	Frequenza Osservata	Frequenza Teorica
0	1362	1362.000
1	405	403.527
2	156	158.909
3	43	41.718
4	9	8.214
5 o oltre	1	1.628
Totale	1976	1976



Risultati del test di buon adattamento con la distribuzione ZIP

Test con distribuzione ZIP		
#V1 per quadrato	Frequenza Osservata	Frequenza Teorica
0	1362	1362.000
1	405	403.527
2	156	158.909
3	43	41.718
4	9	8.214
5 o oltre	1	1.628
Totale	1976	1976

$\lambda = 0.7875997$

$\pi = 0.4300574$

Chi-Squared: 0.2398622

Gradi di libertà: 3

**P-value: 0.971**

Risultati del test di buon adattamento con la distribuzione ZIP

# Risultati e conclusioni





La distribuzione di Poisson sembra descrivere molto bene le frequenze degli impatti, come concluso da Clarke.

**Non c'è alcuna evidenza** per credere che le bombe V1 fossero dotate di un sistema di puntamento straordinariamente avanzato per il periodo.

Grazie per l'attenzione

# Bibliografia

## Articolo:

Clarke, R. D. (1946) An application of the Poisson distribution. *Journal of the Institute of Actuaries*, 72, 481.

The London County Council Bomb Damage Maps 1939-1945' (Amazon link: <https://amzn.to/2Ga6zvT>). Original maps are held at London Metropolitan Archives

## Mappa e dati:

*Map provided by London Metropolitan Archives. Geo-referenced with the support of the National Library of Scotland*

[https://www.google.com/maps/d/u/0/viewer?mid=1VwyxV\\_e\\_LAwzbyJPCAF-C7aCRVNA5W7N&ll=51.4828971341444%2C-0.05419449600869264&z=11](https://www.google.com/maps/d/u/0/viewer?mid=1VwyxV_e_LAwzbyJPCAF-C7aCRVNA5W7N&ll=51.4828971341444%2C-0.05419449600869264&z=11)

<https://www.layersoflondon.org/>

## Librerie:

<https://github.com/geothery/londonShapefiles>

<https://github.com/rstudio/cheatsheets/blob/master/data-visualization-2.1.pdf>

<https://www.rdocumentation.org/packages/vcd/versions/1.4-8/topics/goodfit>