

Chapter 15

Descriptive Statistics with a Purpose

Ronan Conroy
Royal College of Surgeons, Ireland

ABSTRACT

A statistic is a data summary which provides useful information. I use the example of sex work research in Africa to ask the question “useful to whom?”. In over 800 papers on sex work in Africa published over more than two decades, the focus has been on sex workers as spreaders of HIV, enriching the research community while further stigmatising the sex workers who participate in the research. In this study, partly run by sex workers themselves, the focus is on targeting micro-economic aid, and documenting the violent and abusive culture in which sex work is practised. The chapter illustrates the data analysis using Stata. However, it tries to place the descriptive statistics into their wider context in the research agenda, including a discussion of the ethics of research into vulnerable groups.

WHAT IS A STATISTIC?

Sir John Sinclair introduced the word ‘statistic’ into the English language. His explanation of why he did so is still useful to read:

“Many people were at first surprised at my using the words “statistical” and “statistics”, as it was supposed that some in our own language might have expressed the same meaning. [...] the

idea I annex to the term is *an inquiry into the state of a country, for the purpose of ascertaining the quantum of happiness enjoyed by its inhabitants, and the means of its future improvement.*”

Sinclair’s idea of useful facts was not necessarily numbers; any information useful to describing the state of a country and planning for its improvement. Statistics are defined in terms of their usefulness.

A statistic, I propose, is a useful numeric summary of data.

DOI: 10.4018/978-1-60960-097-6.ch015

Now contrast this with the statistics lecturer who breezes into his first class of statistics 101 and says, with an air of enthusiasm that is palpably feigned:

“Suppose we measured the heights of twenty people – how could we summarise what we have found out?”

Now I cannot imagine any reason at all for measuring the heights of twenty people. I cannot imagine a hypothesis that might have driven me to do it, nor a burning question that can be answered with the data. For this reason, there are no statistics you can calculate – there is nothing in the data that answers a question, nothing that can be used to bring about the improvement of people, as Sinclair would have put it. The only question that comes to mind is the mental state of anyone who pops out one fine morning to measure the heights of twenty people.

Not every dataset I analyse is full of excitement, but in each case I cannot start work without understanding what it is that we need to know – what was the purpose of gathering the data. Indeed, I once wrote a paper about the process of finding exactly the way of analysing a very simple dataset that would provide useful information to patient and doctor, showing how only one of the many ways of looking at the data actually gave this information (Conroy, 2002).

I’ve been doing research out in Kenya for almost two decades now. In the course of this, I’ve made a lot of friends, met a lot of people, heard a lot of stories. These somehow shape my idea of what is important, what we need to know, what, in short, of a statistic is.

This chapter takes you through the rationale behind, the conduct and analysis of and my own reflections on the research. The results are published (Elmore-Meegan, Conroy, & Agala, 2004). This is more like the behind-the-scenes documentary.

This is a story about sex workers, some of whom are friends of mine. It is to them I would like to dedicate the chapter.

WHO NEEDS STATISTICS ABOUT SEX WORKERS AND WHY?

Before I introduce some descriptive statistics, I have to explain why we might need them. And to do that, I need to tell you something about sex work and sex work research in sub-Saharan Africa.

The Plague Bringers

Our story starts, as far as I can see, in 1986, with the publication of a paper in the *New England Journal of Medicine* (Kreiss, et al., 1986)

The acquired immunodeficiency syndrome (AIDS) is epidemic in Central Africa. To determine the prevalence of AIDS virus infection in East Africa, we studied 90 female prostitutes, 40 men treated at a clinic for sexually transmitted diseases, and 42 medical personnel in Nairobi, Kenya. Antibody to human T-cell lymphotropic virus Type III (HTLV-III) was detected in the serum of 66 percent of prostitutes of low socioeconomic status, 31 percent of prostitutes of higher socioeconomic status, 8 percent of the clinic patients, and 2 percent of the medical personnel. The presence of the antibody was associated with both immunologic and clinical abnormalities. The mean T-cell helper/suppressor ratio was 0.92 in seropositive prostitutes and 1.82 in seronegative prostitutes (P less than 0.0001). Generalized lymphadenopathy was present in 54 percent of seropositive prostitutes and 10 percent of seronegative prostitutes (P less than 0.0001). No constitutional symptoms, opportunistic infections, or cases of Kaposi’s sarcoma were present. Our results indicate that the epidemic of AIDS virus infection has, unfortunately, spread extensively among urban prostitutes in Nairobi, Kenya. Sexual exposure to men from Central Africa was significantly associated with HTLV-III antibody among prostitutes, suggesting transcontinental spread of the epidemic.

You can see a number of key words in the title: *prostitutes, spread, AIDS, epidemic*. I have

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/descriptive-statistics-purpose/49254

Related Content

Reliability of Performance Factors for Evaluating Electronic Health Information Systems

Margaret Nagwovuma, Gilbert Maiga, Agnes Nakakawaand Emmanuel Eilu (2022). *Building Resilient Healthcare Systems With ICTs* (pp. 153-174).

www.irma-international.org/chapter/reliability-of-performance-factors-for-evaluating-electronic-health-information-systems/298402

An Approach to DNA Sequence Classification Through Machine Learning: DNA Sequencing, K Mer Counting, Thresholding, Sequence Analysis

Sapna Juneja, Annu Dhankhar, Abhinav Junejaand Shivani Bali (2022). *International Journal of Reliable and Quality E-Healthcare* (pp. 1-15).

www.irma-international.org/article/an-approach-to-dna-sequence-classification-through-machine-learning/299963

Enhancing 'Fit' of Health Information Systems Design Through Practice Support

Craig E. Kuziemsky (2010). *Health Information Systems: Concepts, Methodologies, Tools, and Applications* (pp. 642-658).

www.irma-international.org/chapter/enhancing-fit-health-information-systems/49891

Social Orthotics for Youth with ASD to Learn in a Collaborative 3D VLE

James Laffey, Janine Stichterand Matthew Schmidt (2010). *Handbook of Research on Human Cognition and Assistive Technology: Design, Accessibility and Transdisciplinary Perspectives* (pp. 76-95).

www.irma-international.org/chapter/social-orthotics-youth-asd-learn/42829

Tracking of Markers for 2D and 3D Gait Analysis Using Home Video Cameras

Sandro Mihradi, Ferryanto, Tatacipta Dirgantaraand Andi I. Mahyuddin (2013). *International Journal of E-Health and Medical Communications* (pp. 36-52).

www.irma-international.org/article/tracking-of-markers-for-2d-and-3d-gait-analysis-using-home-video-cameras/94632