



1 The Birthday Surprise RWLO

1.1 Introduction

Visit birthday-surprise.blogspot.com for the actual rwlo, or email a short message to bannon@essex.edu if you have any questions about this document or rwlo.

1. RWLO File Format: All files that require mathematical typesetting will be done using $\text{\LaTeX} 2_{\epsilon}$ and then published using Portable Document Format (pdf). These files will be made available on a main blog page, along with ASCII files for data and example code. The blog will allow students to interactively participate from anywhere as long as they have access to a computer connected to the Internet. I will moderate and maintain editorial control of the blog.

2. Eight standard elements of my RWLO

- (a) **Overview:** The Birthday-Surprise is often used in introductory courses on probability and statistics. I believe it is used to show that strange coincidences aren't strange after-all. This rwlo will actually use real data¹ to introduce probabilistic simulation. Having a known distribution of United States births in 1978 will greatly enforce that our assumptions, although false, have little impact on our results. Probability textbooks usually start with the assumption that birth-dates are uniformly distributed and then develop a simple mathematical model to state that only 23 people need to be gathered to have a 50% chance of a birth-date match. So, I'll have the students take actual random samples of the data to see what happens. Here I will emphasize how one should proceed in taking a sample from a population, and the necessity of obtaining a good source of random numbers.

In the event that access to the blog is not feasible, I will provide instructions, worksheets, answer keys, and data to interested faculty.

¹Birth dates, for the year 1978, in the United States.

- (b) **Student Learning Objective:** The students will be able to take a population (United States births in 1978) and develop a scheme to randomly sample twenty-three birthdates. Then the student will be able to ask a simple probabilistic question: is there a birth-date match?
- (c) **Procedure:** Blogging (birthday-surprise.blogspot.com) is to be encouraged where students from diverse regions can share their results. Students will also be able to post questions and comments, this will hopefully create a small community of individuals that will desire to share their ideas. Instructors will only need to give their students the URL (birthday-surprise.blogspot.com) to the blog, and then the instructor can join along to see how the experiment proceeds. It is essential that the instructor monitor their own students to access their development. Interactive birth surveys are integrated within the blog, but students need to use the provided 1978 birth data to do the assignments. If, for some reason, the blog is unavailable, send an email (bannon@essex.edu) message and I will provide electronic copies of the assignments and data.

Summary of Procedure

- i. Students will need to look at a data set from 1978 and answer a series of *simple* statistical questions about this data set. [Worksheet 1]
 - ii. Students will need to serialize the data set. [Worksheet 2]
 - iii. Students will need to sample the serialized data set using a uniform number generator (www.random.org) to determine if there's a birth date match. Or, less preferably, students can use a less sophisticated source of random numbers. [Worksheet 3]
 - iv. Sharing results is encouraged, but the real point is in learning how to take a random sample from a population. Tedious work, but I am hopeful that students will become intrigued with random sampling. For the intellectually inclined, I will provide an introductory text, along with supplements, that will introduce students to computer based stochastic modeling.
- (d) **Content Materials:** The students will need to have access to a Internet connected web browser (if they want participate on the blog), a pdf document reader,² and the ability to read an ASCII file³ of birth data. Analysis of the data can be done by hand, and I will publish several worksheets to guide the students through the analysis. These sheets should be checked by the instructor, especially if the students plan to submit their results to the blog.

²There are many pdf readers, and most are free downloads.

³Any text editor will do.

Electronic copies of all worksheets, answer keys, and data can easily be emailed to interested faculty that are unable, or unwilling to use the blog. All worksheets are available in pdf format for easy distribution. Data sets are in ASCII format for easy electronic manipulation.

- (e) **Assessment:** They will submit their 23 random numbers and the result of their analysis directly to the blog. Instructors should check to see if their sample gives their results. I will publish directions to help facilitate the actual experiment, and, although tedious, sampling a population is simple if you have a good source of random numbers. I need to emphasize here that I have an excellent source of web based random numbers for everyone. However, you need to trust that the students will actually follow the instructions.
- (f) **Course Competencies:** As in introduction to probabilistic modeling. Usually done in introductory probability and statistic courses.
- (g) **Supplementary Resources:** If students have interest, they will be able to download, directly from the blog, a textbook⁴ on mathematical modeling that will give them a firm foundation on how to model probabilistic events using any structured programming language.⁵ This is usually referred to as stochastic modeling or simulation, and it is a very hot field used in many diverse areas. Interested faculty can also request a copy by sending me (bannon@essex.edu) an email message.
- (h) **Recommendations:** I'd really like to see this used as a worldwide distributive learning experiment for motivated high-school students who show promise in mathematics and computer science. Although blogging is encouraged, faculty can still participate by requesting information via email (bannon@essex.edu). The success of this largely depends on how motivated your learners are. Many will 'turn the wheel' and get very little out of this. However, remain forever hopeful that you will capture the attention of the motivated. It is they, that we place hope.

3. Questions can be addressed directly on the blog. Anyone is welcome to post and participate. Although blogging is encouraged, faculty can still participate by requesting information via email (bannon@essex.edu).

⁴A draft version of *Deterministic Uncertainties* authored by Ron Bannon.

⁵C++ is preferred and is extensively used because it's readily available, and is considered robust for probabilistic modeling.