Probability Theory and Statistics Inference

Lesson 1

Contents

Sample Espace

- Experiment: a repeatable procedure with well-defined possible outcomes.
- Examples:
    - Toss the coin, report if it lands heads or tails.
    - Toss the coin 3 times, list the results.
    - Measure the mass of a proton.
    - Count the number of taxis that pass for and avenue someday.

Sample Espace

- Set of all possible outcomes. We usually denote the sample space by Ω.
- Examples:
  - $\Omega = \{H, T\}$
  - $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
  - $\Omega = [0, \infty)$
  - $\Omega = \{0, 1, 2, 3, 4, \ldots\}$.

Events

- A subset of the sample space.
- Let $\Omega = \{ \omega_1, \omega_2, \ldots, \omega_n \}$ be a sample space, then:
  - Impossible event: $\{\}$
  - Unitary event: $\{ \omega_i \}$
  - Safe event: Ω
- Satisfies all set theory.

Probability function

- For a sample space Ω a probability function P assigns to each outcome ω a number $P(\omega)$ called the probability of ω.

  $\Omega = \{ \omega_1 , \omega_2 , \ldots , \omega_n \}$

  - $0 \leq P(\omega) \leq 1$
  - The sum of the probabilities of all possible outcomes is 1

$$\sum_{j=1}^{n} P(\omega_j) = 1$$

Some rules of Probability

- For events A, L and R contained in a sample space Ω.
-
  - $P(A') = 1 - P(A)$.
  - If L and R are disjoint then: $P(L \cup R) = P(L) + P(R)$
  - If L and R are not disjoint, we have the inclusion-exclusion principle:

$$P(L \cup R) = P(L) + P(R) - P(L \cap R)$$

$$P(A) = \frac{\text{Number or favorable outcomes to A}}{\text{Total number of outcomes}}$$

Examples and Code

Example Consider the 2 events, A: 'X is a multiple of 2'; B: 'X is odd and less than 10'. Suppose $P(A) = .6$ and $P(B) = .25$.
(i) What is $A \cap B$?
(ii) What is the probability of $A \cup B$?

Example Let A, B and C be the events X is a multiple of 2, 3 and 6 respectively. If $P(A) = .6$, $P(B) = .3$ and $P(C) = .2$ what is $P(A \text{ or } B)$?

EJEMPLO

La demanda de dos productos A y B varía aleatoriamente en un rango de 1000 a 5000 kilogramos. El distribuidor decide bajar el precio de venta de ambos productos si la suma de sus demandas varía de 3000 a 5000 Kg. Calcular la probabilidad de que el precio de venta de ambos productos baje.

Examples and Code

Problem 3. (20 pts.) **Birthdays: counting and simulation.**
Ignoring leap days, the days of the year can be numbered 1 to 365. Assume that birthdays are equally likely to fall on any day of the year. Consider a group of n people, of which you are not a member. An element of the sample space Ω will be a sequence of n birthdays (one for each person).

(a) Define the probability function P for Ω..

(b) Consider the following events:
A: "someone in the group shares *your* birthday"
B: "some two people in the group share a birthday"
C: "some three people in the group share a birthday"

Carefully describe the subset of Ω that corresponds to each event.

(c) Find an exact formula for $P(A)$. What is the smallest n such that $P(A) > .5$?

(d) Justify why n is greater than $\frac{365}{2}$ without doing any computation. (We are looking for a short answer giving a heuristic sense of why this is so.)

(e) Use R simulation to estimate the smallest n for which $P(B) > .9$.

(f) Find an exact formula for $P(B)$.

Examples and Code

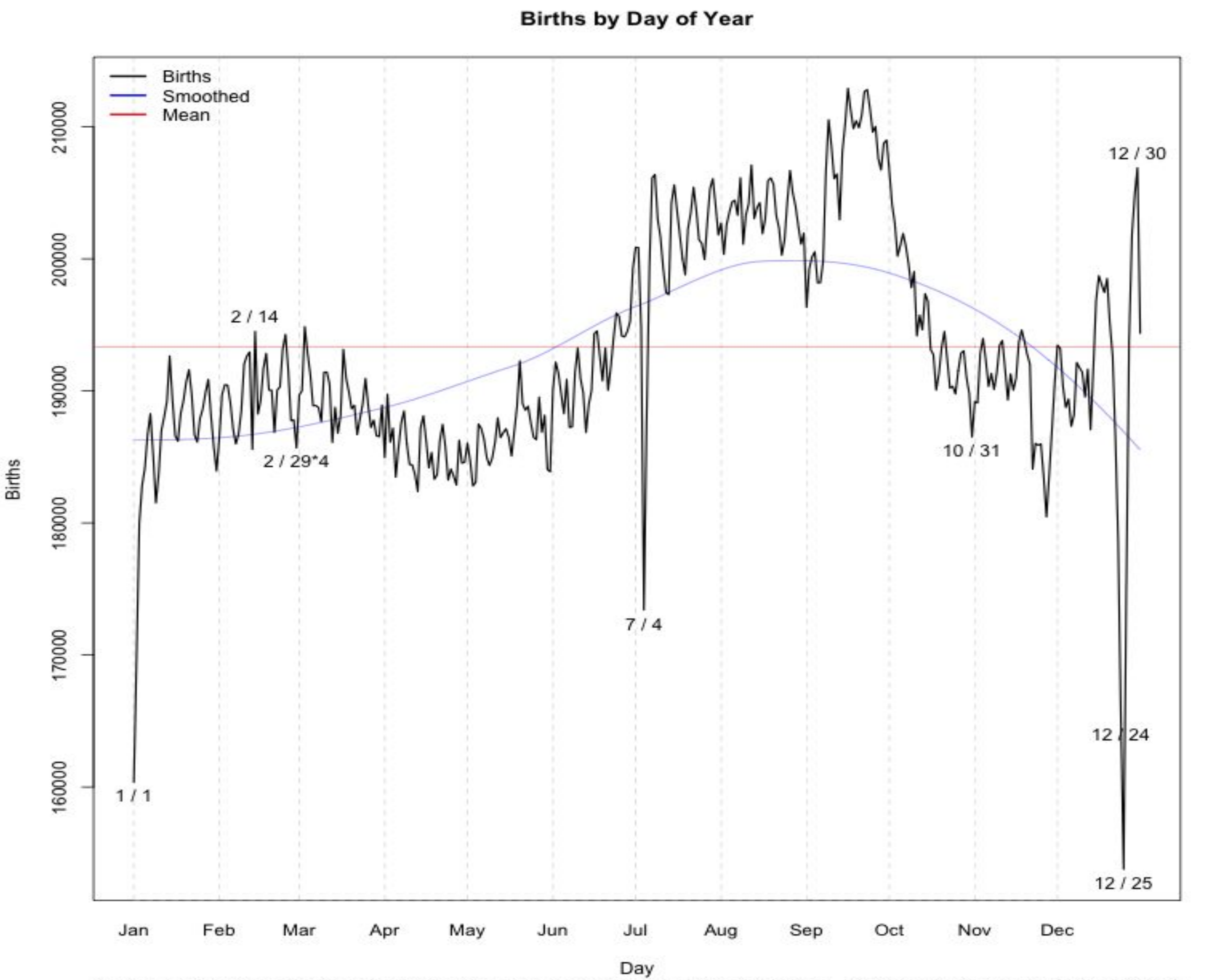
```
import sys

# main params
n = int(sys.argv[1])

def prob_B():
	total_days = 365
	decrease_days = 365
	output = 1

	for i in range(n):
		a = decrease_days / (total_days)
		decrease_days -= 1
		output *= a

	return (1.0 - output)
```



Source: National Vital Statistics System natality data, as provided by Google BigQuery. Graph by Chris Mulligan (chmullig.com)

Math Team

Christian Córdova

Inserta tu texto aquí Inserta tu texto aquí Inserta tu texto aquí Inserta tu texto aquí

José Castro

Inserta tu texto aquí Inserta tu texto aquí Inserta tu texto aquí Inserta tu texto aquí


Statistics
Density
population
Probability
experiment
intersection
Venn-Diagram
probability-tree
box-plot
observation
distribution
Standard
standard
randomness
histogram
bargraph
event
z-score
Median
union
Mode
Normal
independence
Mean
dotplot
pie-chart
conditional
Deviation
Random
chance
Skewed
stem-leaf
Sample