U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Capítulo 7

Introdução à Inferência Estatística

AMG, JFO (v8 – 2017)
adaptado de: *Estatística*,
Rui Campos Guimarães,
José A. Sarsfield Cabral

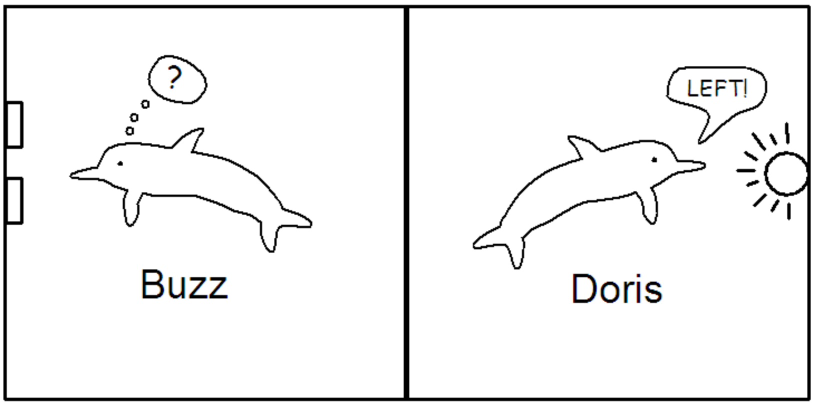
Slide 7a.-1

Conteúdo

Slide 7a.0

7.1 Case Study: Can Dolphins Communicate?

- Caso de estudo retirado do livro "Introduction to Statistical Investigations", Nathan Tintle, *et al.*, Wiley



- Ver actividade disponível no Moodle.

Slide 7a.1

7.2 Case Study: Real vs Fake Coin Flips

In these two sequences of 200 coin flips, 1 stands for a head and 0 for a tail. One of the sequences shows what happened when a student actually tossed a fair coin 200 times. The other sequence was made up by a student trying to fake flipping a coin 200 times. Which do you think is the real sequence and which is the fake?

- 01000 01101 01000 00110 10111 11111 10101 01110 11100 11010 11111 01101 10010 10100 00001 10101 01000 10000 11000 01000 01000 01110 10110 00100 01001 10111 01111 10111 00111 01100 00000 10001 00010 01000 10011 11011 10001 00000 01101 00011
- 01010 00100 00110 10001 00111 01101 01010 00011 00110 01110 11100 11100 01110 11110 00101 01111 01001 11011 01000 11011 10111 10010 11011 10001 01110 11000 11100 00111 01011 11010 01100 00101 00100 10110 01000 10000 11110 10111 00110 11011

Slide 7a.2

- After a little discussion, the students almost unanimously agree that the *second sequence* looks like the one made from flipping a coin.
- They say such things as, "*The heads and tails look more mixed up in the second sequence*" and "*The first sequence sometimes has too many heads in a row and too many tails in a row*."
- How to make a more educated decision as to which sequence is the real one?
- One student immediately replies, "*Try it with a real coin. Flip a real coin 200 times!*"
- How the new sequence of real coin flips would tell them which of the two original sequences is the real one?
- *Silence.*
- The students decide to discuss this question in small groups.

Slide 7a.3

- Percentage of heads is not in doubt. Both sequences have about the right number of heads and tails.

- However, *the first sequence has streaks of heads and tails that seem too long*. Specifically, the first sequence has nine heads in a row and the students think that should not happen.
- Why?
- Suggestion: flip a coin 200 times and counting the longest streak of heads.
- *It is necessary to do this many times* and see *how often they get a streak of nine heads or more*.
- How?
  - flip a coin 200 times or
  - use a computer to simulate flipping a coin 200 times.
- Each student counts the longest run of heads in his or her sequence.
- The students make a histogram of their results.

Slide 7a.4

- It is shaped approximately like the following theoretical (non-normal) sampling distribution:

Longest Run	*Frequency*	
4	3	⋆⋆⋆
5	17	⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆
6	26	⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆
7	22	⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆
8	14	⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆⋆
9	9	⋆⋆⋆⋆⋆⋆⋆⋆⋆
10	5	⋆⋆⋆⋆⋆
11	3	⋆⋆⋆
12	1	⋆

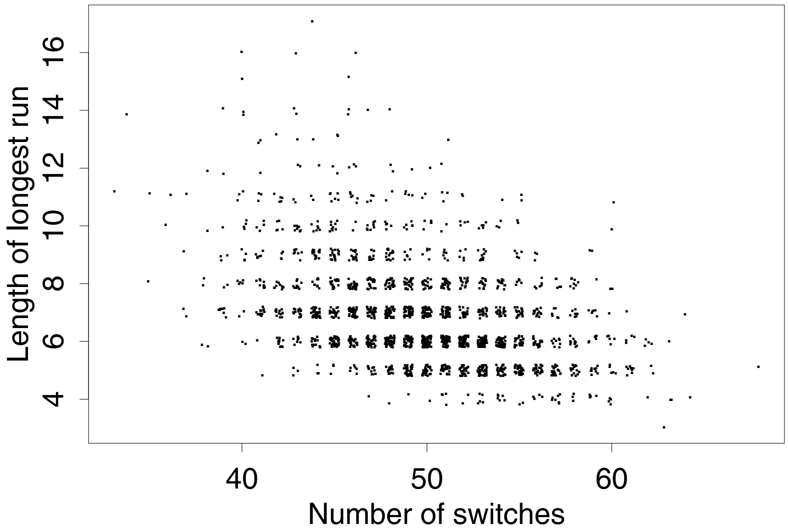
- The students now realize that *the first sequence is more likely to be the real one*. A longest run of heads of four, which the second sequence has, is fairly unlikely to occur in 200 tosses of a fair coin. A longest run of heads of nine, which the first sequence has, is much more likely.

Slide 7a.5

- To assess if the students got the point of the lesson and understand the mechanics of the histogram:
  - If a basketball player is a 50% field goal shooter, and shoots 200 times in a series of games, what would you expect the longest streak of baskets to be if the player does not exhibit unusually streaky behavior? How could you tell if the player was "hot"?
  - What could be some of the reasons why people tend to believe that players exhibit "streaks" and "slumps" in games like basketball?
  - Would a histogram of the lengths of the longest run of tails be the same as or different from that of the longest run of heads? Explain.
  - In this activity you constructed a histogram of the longest run of heads in a sequence of 200 coin flips. The probability of a head was 0.5 on each flip. Suppose instead you were constructing a histogram of the longest run of heads in a sequence of 200 coin flips for a weighted coin for which the probability of a head was 0.6 on each flip. How would this new histogram be different from the one you constructed? How would it be the same?

Slide 7a.6

- Is there other statistic that could be used?



Slide 7a.7

7.3 Statistical Inference

With the Dolphins case study, we saw how to say things like:

- We have strong evidence that the long-run probability Buzz pushes the correct button is larger than 0.5 with a curtain wall.
- We do not have strong evidence that the long-run probability Buzz pushes the correct button is larger than 0.5 with a wood wall.

⇒ How to to identity what is *statistically significative* (i.e., distinguish if it is due to chance or not)

(topic to be studied in detail in the chapter "Hypothesis Tests")

We also want a method that says:

- "I'm confident that Buzz pushes the correct button 91 to 96% of times with a curtain wall."

Slide 7a.8

Estimation allows us:

- to obtain an estimative of a population parameter (*Point Estimation*), and
- to know how large the effect is, through an interval of values (*Interval Estimation*).

Examples

- 93.75% is an estimative of the "true" (long-run) probability of the Buzz ability to push the correct button with a curtain wall.

- We can be 95% confident that the "true" effect of taking bi-daily aspirin will reduce the rate of heart attacks somewhere between 30% and 50%.
- "If the election were held today, would you vote for Barack Obama or Mitt Romney?" 51% responded Obama (margin of error is ± 3 percentage points).

Slide 7a.9

$\Rightarrow$ These interval estimates of a population parameter are called confidence intervals.

There are two main strategies to obtain them:

- Using simulation to obtain the standard deviation of the null distribution to determine the width of the interval.
- Through traditional theory-based methods (based on the Central Limit Theorem).

$\Rightarrow$ We'll only study the traditional theory-based methods

Slide 7a.10

We'll see how to find good "point estimates", "confidence intervals" and perform hypothesis tests for the usual population parameters, including:

- a population mean μ
- the difference in two population means, $\mu_1 - \mu_2$
- a population variance σ^2
- the ratio of two population variances, $\frac{\sigma_1^2}{\sigma_2^2}$
- a population proportion p
- the difference in two population proportions, $p_1 - p_2$

We will work on not only obtaining formulas for the estimates and intervals, but also on arguing that they are "good" in some way (unbiased, for example). We'll also address practical matters, such as how sample size affects the length of our derived confidence intervals. Slide 7a.11