

## EXERCISE II: PROBABILITY

### TOPICS COVERED:

- A, B,  $A^C$ ,  $A \cap B$ ,  $A \cup B$
- Disjoint and independent
- Conditional probability

- REQUIRED DATASET: IPUMS-International**

- REQUIRED VARIABLES:**

1. COUNTRY
2. YEAR
3. INTERNET (household has internet access)
4. OWNERSHIP (ownership of dwelling)

*[The only preselected variables that are needed in this exercise are COUNTRY and YEAR. Make sure to remove all of the other preselected variables by unchecking the blue boxes next to them. This will reduce the size of your data file and also make it easier to view the data in R.]*

- RECOMMENDED SAMPLES:**

1. Brazil 2010\*\*\*

\*\*\*Limit sample to 10,000 households.

---

**Let A = households from Brazil who have internet access**

**Let B = people from Brazil who have ownership of their dwelling**

---

### ❖ Question 1

**A. What population was asked about internet access [[INTERNET](#)] in the Brazil 2010 census? Should you include or exclude missing values [0: NIU (not in universe)] from your calculations?**

Persons living in permanent private housing units with a computer were asked about internet access. Persons in the NIU (not in universe) category were not asked about internet access and should be excluded from calculations.

**B. What is the probability that a randomly selected person in Brazil lives in a household (with a computer) that has internet access? ( $P(A)$ )?**

$$P(A) = 8435 / 11022 = 0.7652876$$

**C. What is the probability that a randomly selected person in Brazil lives in a household (with a computer) that doesn't have internet access? ( $P(A^c)$ )?**

$$P(A^c) = 2587 / 11022 = 0.2347124$$

---

❖ Question 2

**A. Were the same people asked about internet access and dwelling ownership [[OWNERSHIP](#)] in the Brazil 2010 census?**

Only persons living in (occupied permanent private) households with a computer were asked about internet access. All person living in occupied permanent private households were asked about dwelling ownership.

**B. What is the probability that a randomly selected person from Brazil has ownership of their dwelling? ( $P(B)$ ) [exclude 0:NIU]**

$$P(B) = 25157 / 33186 = 0.7580606$$

**C. What is the probability that a randomly selected person from Brazil does not have ownership of their dwelling? ( $P(B^c)$ )?**

$$P(B^c) = 8029 / 33186 = 0.2419394$$

---

❖ Question 3

**Are the variables INTERNET and OWNERSHIP disjoint (mutually exclusive)? Why/why not?**

No, because both events can happen at the same time. It is possible for a household to have internet access and to have ownership of their dwelling.

---

#### ❖ Question 4

Are the variables INTERNET and OWNERSHIP independent? If they are independent, then  $P(B|A) = P(B)$  and  $P(A|B) = P(A)$ . How should you restrict the sample for this test to be sure you are calculating probabilities for the same population for both variables?

The sample should be restricted to only persons that were asked about internet access. Persons with value 0: NIU for INTERNET should be excluded.

$$P(B|A) = 6648 / 8435 = 0.7881446$$

$$P(B) = 8515 / 11022 = 0.7725458$$

$$P(A|B) = 6648 / 8515 = 0.7807399$$

$$P(A) = 8435 / 11022 = 0.7652876$$

Both formulas are true in this case as the probabilities are very similar, so this means the probability that a person lives in a household with internet access is not dependent upon whether a person has ownership of their dwelling.

---

#### ❖ Question 5

- A. What is the probability that a random person from Brazil has internet access and owns their dwelling? ( $P(A \cap B)$ ?)

$$P(A \cap B) = P(A) * P(B) = 0.7652876 * 0.7725458 = 0.5912197$$

- B. What is the probability that a random person from Brazil has internet access and does not own their dwelling? ( $P(A \cap B^c)$ ?)

$$P(A \cap B^c) = P(A) * P(B^c) = 0.7652876 * 0.2419394 = 0.1851532$$

- C. What is the probability that a random person from Brazil does not have internet access and owns their dwelling? ( $P(A^c \cap B)$ ?)

$$P(A^c \cap B) = P(A^c) * P(B) = 0.2347124 * 0.7725458 = 0.1813267$$

**D. What is the probability that a random person from Brazil does not have internet access and does not own their dwelling? ( $P(A^c \cap B^c)$ )?**

$$P(A^c \cap B^c) = P(A^c) * P(B^c) = 0.2347124 * 0.2419394 = 0.0567862$$

---

❖ Question 6

**A. What is the probability that a randomly selected person from Brazil has internet access or owns their dwelling? ( $P(A \cup B)$ )?**

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) = 0.7652876 + 0.7725458 - 0.5912197 \\ &= 0.9466137 \end{aligned}$$

**B. What is the probability that a randomly selected person from Brazil has internet access or does not own their dwelling? ( $P(A \cup B^c)$ )?**

$$\begin{aligned} P(A \cup B^c) &= P(A) + P(B^c) - P(A \cap B^c) = 0.7652876 + 0.2419394 - 0.185153 \\ &= 0.822074 \end{aligned}$$

**C. What is the probability that a randomly selected person from Brazil does not have internet access or owns their dwelling? ( $P(A^c \cup B)$ )?**

$$\begin{aligned} P(A^c \cup B) &= P(A^c) + P(B) - P(A^c \cap B) = 0.2347124 + 0.7725458 - \\ &0.1813267 = 0.8259315 \end{aligned}$$

**D. What is the probability that a randomly selected person from Brazil does not have internet access or does not own their dwelling? ( $P(A^c \cup B^c)$ )?**

$$\begin{aligned} P(A^c \cup B^c) &= P(A^c) + P(B^c) - P(A^c \cap B^c) = 0.2347124 + 0.2419394 - \\ &0.0567862 = 0.4198656 \end{aligned}$$