STATISTICS

Module 2 *of*
MATHEMATICS AND STATISTICS

Marina Santacroce

marina.santacroce@unibocconi.it
Department of Decision Sciences, Bocconi University

L5. Conditional probability and independence. Applications and exercises.

Outline

Conditional Probabilities

Independence

- Independent events
- Bernoulli trials
- Conditional Independence (*)

Exercises

Conditional Probabilities

Def. Consider two events E and F with $\mathbb{P}(F) > 0$. The conditional probability of E given F is

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Thus,

$$\mathbb{P}(E \cap F) = \mathbb{P}(F)\mathbb{P}(E|F).$$

For n events $\Rightarrow$ the Multiplication Rule:

If $\mathbb{P}(E_1 \cap E_2 \cap \ldots E_{n-1}) > 0$

$$\mathbb{P}(E_1 \cap E_2 \cap \ldots E_n) = \mathbb{P}(E_1)\,\mathbb{P}(E_2|E_1)\,\mathbb{P}(E_3|E_1 \cap E_2) \cdots \mathbb{P}(E_n|E_1 \cap E_2 \ldots E_{n-1})$$

Independent events

Def. We say E and F are *independent* if the occurrence of one does not modify the probability of the other, i.e. if $\mathbb{P}(F) > 0$, E and F are independent if

$$\mathbb{P}(E|F) = \mathbb{P}(E).$$

Using the definition of conditional probability, we immediately have an alternative definition of independence.

Def.($\star$) We say E and F are *independent* if

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F). \qquad (\star)$$

$\Rightarrow$ independence is symmetric!

$\Rightarrow$ the definition ($\star$) includes the case of zero-probability (or null) events.

- ◇ Show that if E and F are independent, E^c and F^c are independent (and also E and F^c).
- ◇ Show that if $E \subset F$ the events are independent if and only if E is a null event or F is a certain event ($\mathbb{P}(E) = 0$ or $\mathbb{P}(F) = 1$).

Def. The events E, F and G are said independent if

$$\begin{aligned}
\mathbb{P}(E \cap F \cap G) &= \mathbb{P}(E)\mathbb{P}(F)\mathbb{P}(G)\\
\mathbb{P}(E \cap F) &= \mathbb{P}(E)\mathbb{P}(F)\\
\mathbb{P}(E \cap G) &= \mathbb{P}(E)\mathbb{P}(G)\\
\mathbb{P}(G \cap F) &= \mathbb{P}(G)\mathbb{P}(F)
\end{aligned}$$

NOTE

$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F) \qquad \mathbb{P}(E \cap G) = \mathbb{P}(E)\mathbb{P}(G) \qquad \mathbb{P}(G \cap F) = \mathbb{P}(G)\mathbb{P}(F)$$

are not *sufficient* for three events to be independent!

Example: An urn contains 8 labeled from 1 to 8 balls. A ball is drawn. Are $E = \{$the number is 1,2,3 or 4$\}$, $F = \{$the number is 3,4,5 or 6$\}$ and $G = \{$the number is 1,2,5 or 6$\}$ independent?

Consider n events $E_1, \dots E_n$.

Def. We say $E_1, E_2, \dots E_n$ are independent if for any $k \in \{2, \dots n\}$ and any choice $\{i_1, \dots i_k\}$ one has

$$\mathbb{P}(E_{i_1} \cap E_{i_2} \cap \dots E_{i_k}) = \prod_{i_j \in \{i_1 \dots i_k\}} \mathbb{P}(E_{i_j}).$$

Independent Bernoulli trials

Example: An infinite sequence of independent trials is to be performed. Each trial is a success with probability p and a failure with probability $1 - p$. What is the probability that

1) at least one success occurs in the first n trials;
2) exactly k successes occur in the first n trials;
3) (*) all trials are successes?

Conditional independence (*)

Def. Consider F with $\mathbb{P}(F) > 0$. We say $E_1, E_2, \ldots E_n$ are conditionally independent given F if for any $k \in \{2, \ldots n\}$ and any choice $\{i_1, \ldots i_k\}$ one has

$$\mathbb{P}(E_{i_1} \cap E_{i_2} \cap \ldots E_{i_k} | F) = \prod_{i_j \in \{i_1 \ldots i_k\}} \mathbb{P}(E_{i_j} | F).$$ [1]

Example: Data from joint United Nation Programme on HIV/AIDS (2006): We have a test such that $\mathbb{P}(T^+ | \mathrm{HIV}^+) = 0.99$ (sensitivity) and $\mathbb{P}(T^- | \mathrm{HIV}^-) = 0.99$ (specificity). Suppose the prevalence of HIV in a country is 1%. Given the first test is positive, event T_1^+, what is the probability of being HIV positive? The policy for a positive HIV test is a follow-up confirmatory test. Given the 2nd test is positive, event T_2^+, what is the probability of being HIV positive?
Suppose that the events T_1^+ and T_2^+ are conditionally independent given HIV^+ (HIV^-).

[1] i.e. if are independent with respect to the conditional probability $\mathbb{P}(\cdot | F)$.

Exercises

1. In answering a question on a multiple-choice test, a student either knows the answer or guesses. Let p be the probability the student knows the answer and assume that, if he does not know it, his answer will be correct with probability $\frac{1}{m}$, where m is the number of choices. What is the conditional probability that a student knew the answer given he answered correctly?
2. Suppose you have three coins: one regular coin, another one with two heads on its sides and the latter with two tails. You choose one randomly and you flip it. Given the flip lands on head, what is the probability the other side of the coin is head?
3. A bin contains 3 different flashlights. The probability that a type 1 flashlight will give over 1000 hours of use is 0.7, with the corresponding probabilities for type 2 and 3 flashlights being 0.4 and 0.3, respectively. Suppose that 20% of the flashlights are of type 1, 30% of type 2. What is the probability that a randomly chosen flashlight will give more than 1000 hours of light? Given it did, what is the probability that it was a type j, with $j : 1, 2, 3$?
4. (**Coupons**) There are n types of coupons, and each new one collected is independently of type i with probability p_i, with $i : 1, 2, \dots, n$.
   Suppose k coupons are to be collected. What is the probability that A_i={there is at least one of type i among those collected}?
   Find $\mathbb{P}(A_i \cup A_j)$ and $\mathbb{P}(A_i | A_j)$.

5. (**Urns**, with replacement) Three urns have the following composition:
Urn 1: has 4 white, 2 red and 2 green balls;
Urn 2: has 2 white, 4 red and 2 green balls;
Urn 3: has 3 white, 3 red and 2 green balls.

   You first select an urn at random and then you draw two balls from the urn with replacement.[2] Compute:

   a) the probability of drawing 2 red balls.
   b) the probability of having chosen urn 2 given 2 red balls have been drawn.
   c) the probability of drawing 2 balls of different colors.

6. (**Urns**, without replacement) Repeat the previous exercise in the case without replacement.

7. (**Series and Parallel Systems**)
A system composed of n separate components. The system is said to be in parallel if it functions when at least one of its components functions. While the system is said to be in series if it functions when all its components function. Suppose component i, independently of the others, functions with probability p_i, $i : 1, \dots, n$. Find the probability that the system is functional (in both cases).

[2] before selecting the second ball you replace the first ball in the urn.