Independent events; Random variables

For two events E and F we say that they are independent if:

$$P(E\cap F) = P(E)\cdot P(F).$$

Assuming that $P(F)>0$, looking at $P(E|F)$, we get that if E and F are independent:

$$P(E|F) = \frac{P(E\cap F)}{P(F)} = \frac{P(E)\cdot P(F)}{P(F)} = P(E),$$

informally, the occurrence of one event does not affect the probability of occurrence of the other event.

P1 Suppose we toss 2 fair dice. Let S_6 be the event that the sum of the dice is 6, and S_7 that the sum is 7. Let E_4 be the event that the first die is 4. Are S_6 and E_4 independent? What about S_7 and E_4.

Sol: We can represent the sample space for the experiment as

$$\Omega = \{(i,j) : i,j\in\{1,2,\dots,6\}\};$$

where the first coordinate represents the number on the first die and the second on the second die.

$$S_6 = \{(1,5),(2,4),(3,3),(4,2),(5,1)\};$$

$S_7 = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\};$

$E_4 = \{(4,1),(4,2),\dots,(4,6)\}.$

Since the dice are fair, we have:

$$P(S_6) = \frac{\text{number of elements in } S_6}{\text{number of elements in } \Omega} = \frac{5}{36} \quad ;$$

$$P(S_7) = \frac{6}{36} = \frac{1}{6} \quad \text{and} \quad P(E_4) = \frac{6}{36} = \frac{1}{6}.$$

Now, $S_6 \cap E_4 = \{(4,2)\}$ and $S_7 \cap E_4 = \{(4,3)\}$, and so:

$$P(S_6 \cap E_4) = \frac{1}{36} \neq \frac{5}{36} \cdot \frac{1}{6} = P(S_6) \cdot P(E_4),$$

hence S_6 and E_4 are not independent, but

$$P(S_7 \cap E_4) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(S_7) \cdot P(E_4), \text{ and}$$

so S_7 and E_4 are independent events.

P2 Show that if E and F are independent, then E and F^c are independent as well.

Sol: We need to prove that $P(E \cap F^c) = P(E) \cdot P(F^c)$. Since $E \cap F^c = E \setminus F$, we have:

$P(E \cap F^c) = P(E \setminus F) = P(E) - P(E \cap F)$, this is true because $E \setminus F$ and $E \cap F$ are mutually exclusive and

$(E \setminus F) \cup (E \cap F) = E$ and so $P(E \setminus F) + P(E \cap F) = P(E)$.
Now, from the independence of E and F: $P(E \cap F) = P(E) \cdot P(F)$, and so:

$$P(E) - P(E \cap F) = P(E) - P(E) \cdot P(F) = P(E) \cdot (1 - P(F))$$
$$= P(E) \cdot P(F^c) \text{, hence}$$

$$P(E \cap F^c) = P(E) \cdot P(F^c) \text{, i.e. } E \text{ and } F^c \text{ are independent.}$$

P3 Independent trials of rolling a pair of fair dice are performed. What is the probability that an outcome of 5 appears before an outcome of 7, when the outcome of a roll is the sum of the dice.

Sol: Let E_n be the event that in the first $n-1$ rolls there were no 5 or 7 and on the n-th roll 5 is obtained.

In P1 we obtained $P(\text{sum} = 7 \text{ (for one roll)}) = \frac{6}{36} = \frac{1}{6}$

and similarly $P(\text{sum} = 5) = \frac{4}{36} = \frac{1}{9}$, since rolling 5 or 7 (as sum) are mutually exclusive events,

$$P(\text{rolling a sum different from } 5 \text{ or } 7) = 1 - P(\text{rolling } 5) - P(\text{rolling } 7)$$
$$= 1 - \frac{4}{36} - \frac{6}{36} = \frac{26}{36} = \frac{13}{18}.$$

Throws are independent, hence:

$$P(E_n) = \underbrace{\frac{13}{18} \cdot \ldots \cdot \frac{13}{18}}_{n-1 \text{ times}} \cdot \frac{4}{36} = \frac{1}{9} \left(\frac{13}{18}\right)^{n-1}$$

Notice that $\bigcup_{n=1}^{+\infty} E_n$ is the event that a sum of 5 is obtained before a sum of 7; and since E_n's are mutually exclusive:

$$P\left(\bigcup_{n=1}^{+\infty} E_n\right) = \sum_{n=1}^{+\infty} P(E_n) = \frac{1}{9} \cdot \underbrace{\sum_{n=1}^{+\infty} \left(\frac{13}{18}\right)^{n-1}}_{\text{a geometric series}} = \frac{1}{9} \cdot \frac{1}{1-\frac{13}{18}}$$

$$= \frac{1}{9} \cdot \frac{1}{\frac{5}{18}} = \frac{1}{9} \cdot \frac{18}{5} = \frac{2}{5}.$$

P4 (The gambler's ruin problem)

Two gamblers, A and B, bet on the outcomes of successive flips of a coin. On each flip, if it's heads A takes 1 euro from B, and if it's tails B takes 1 euro from A. They play until one of them runs out of money. Assume that successive flips are independent, and that probability of obtaining heads is p. What is the probability that A ends up with all the money if he starts with k euros and B with $N-k$ euros?

Sol: Let E be the event that A ends up with all the money. Let $P_k = P(E)$ be the probability of E when A starts with k euros. Let H be the event that the first flip lands on heads; then

$$\begin{aligned} P_k = P(E) &= P(E|H) \cdot P(H) + P(E|H^c) \cdot P(H^c) \\ &= P_{k+1} \cdot p + P_{k-1} \cdot (1-p) ; \end{aligned}$$

So: $P_k = p \cdot P_{k+1} + (1-p) \cdot P_{k-1} \Rightarrow$

$$p\left(P_{k+1} - P_k\right) = (1-p)\left(P_k - P_{k-1}\right) \quad \Big| : p$$

$$P_{k+1} - P_k = \frac{1-p}{p}\left(P_k - P_{k-1}\right).$$

Denoting $\frac{1-p}{p}$ by t, and using the fact that $P_0 = 0$,

$$P_2 - P_1 = t \cdot (P_1 - P_0) = t \cdot P_1$$

$$P_3 - P_2 = t \cdot (P_2 - P_1) = t^2 \cdot P_1$$

$$\vdots$$

$$P_N - P_{N-1} = t \cdot (P_{N-1} - P_{N-2}) = t^{N-1} \cdot P_1$$

Adding the first $k-1$ equations together we obtain:

$$P_k - P_1 = \left(t + t^2 + \dots + t^{k-1}\right) P_1 \Rightarrow P_k = \left(1 + t + \dots + t^{k-1}\right) P_1, \text{ so}$$

$$P_k = \begin{cases} \frac{1-t^k}{1-t} P_1 & \text{if } t \neq 1 \text{ i.e. } p \neq \frac{1}{2} \\ k P_1 & \text{if } t = 1 \text{ i.e. } p = \frac{1}{2} \end{cases}$$

Since $P_N = 1$, we get $1 = \frac{1-t^N}{1-t} P_1 \Rightarrow P_1 = \frac{1-t}{1-t^N}$

if $t \neq 1$ and $1 = N \cdot P_1 \Rightarrow P_1 = \frac{1}{N}$ if $t = 1$, so

$$P_k = \begin{cases} \frac{1-t^k}{1-t^N} & \text{if } p \neq \frac{1}{2} \\ \frac{k}{N} & \text{if } p = \frac{1}{2} \end{cases}.$$

A random variable is a function from a sample space Ω to the set of real numbers. Usually we denote random variables by letters X, Y, Z, etc.

Notation: In probability when we have a random variable $X: \Omega \to \mathbb{R}$ and a real number r, we write $P(X=r)$ instead of more cumbersome, but more precise, $P(\{\sigma \in \Omega : X(\sigma) = r\})$, and similary for $P(X \geq r)$, $P(X \leq r)$, $P(r_1 \leq X \leq r_2)$.

P5 Three balls are randomly selected from an urn containing 20 balls numbered 1 through 20. We bet that at least one of the balls that we draw has a number as large or larger than 17. What is the probability that we will win the bet?

$\boxed{\text{Sol:}}$ Our sample space for the experiment is:

$$\Omega = \{ \{i,j,k\} : i,j,k \in \{1,2,\dots,20\} \text{ and } i \neq j \neq k \}$$

The number of elements in Ω is $\binom{20}{3}$. Let $X: \Omega \to \mathbb{R}$ be the random variable defined by:

$$X(\{i,j,k\}) = \max(\{i,j,k\}), \quad \forall \{i,j,k\} \in \Omega,$$

i.e. X maps outcome to the largest number on the selected balls. We want to find out what $P(X \geq 17)$ is.

Take $n \in \{3, 4, \ldots, 20\}$, then:

$$P(X=n) = P(\{\sigma \in \Omega : X(\sigma) = n\}), \text{ and note that}$$

$\{\sigma \in \Omega : X(\sigma) = n\}$ is the event that the largest number is n. Since n is fixed, we are free to choose 2 number less than n to form a set σ such that $X(\sigma) = n$. This means that we have $\binom{n-1}{2}$ elements in our set $\{X = n\}$, and hence:

$$P(X=n) = \frac{\binom{n-1}{2}}{\binom{20}{3}}, \quad \forall n \in \{3, \ldots, 20\}.$$

From this we compute:

$$P(X \geq 17) \overset{\text{axiom 3. (mutually exclusive events)}}{=} P(X=17) + P(X=18) + P(X=19) + P(X=20)$$

$$= \frac{1}{\binom{20}{3}} \left(\binom{16}{2} + \binom{17}{2} + \binom{18}{2} + \binom{19}{2} \right)$$

$$= 0.150 + 0.134 + 0.119 + 0.105 = 0.508 \, .$$

For a random variable X the function $F_X : \mathbb{R} \to \mathbb{R}$ defined by:

$$F_X(r) = P(X \leq r), \quad \forall r \in \mathbb{R}$$

is called **distribution function** of X, and **probability mass (density) function** p of X is defined by:

$$p(r) = P(X = r). \quad \left(\left(\text{thanks for discrete random variables}\right)\right)$$

P6 The probability mass function of a random variable X is given by $p(i) = c\lambda^i / i!$, for $i = 0, 1, 2, \dots$; where λ is some positive value.
Find $P(X=0)$ and $P(X>2)$!

Sol: From $\sum_{i=0}^{+\infty} \frac{c\lambda^i}{i!} = \sum_{i=0}^{+\infty} p(i) = 1$ we have:

$$1 = c \cdot \sum_{i=0}^{+\infty} \frac{\lambda^i}{i!} = c \cdot e^{\lambda} \Rightarrow \quad c = e^{-\lambda}; \text{ so}$$

$$p(i) = e^{-\lambda} \cdot \lambda^i / i! \quad \forall i = 0, 1, 2, \dots$$

From this: $P(X=0) = e^{-\lambda} \cdot \lambda^0 / 0! = e^{-\lambda}$, and

$$P(X>2) = 1 - P(X \leq 2) = 1 - P(X=2) - P(X=1) - P(X=0)$$
$$= 1 - e^{-\lambda} \cdot \lambda^2/2 - e^{-\lambda} \cdot \lambda - e^{-\lambda}.$$

P7 The probability mass function of a random variable X is given by:

$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = \frac{1}{8}, \quad p(4) = \frac{1}{8}.$$

Find its distribution function F and draw its graph.

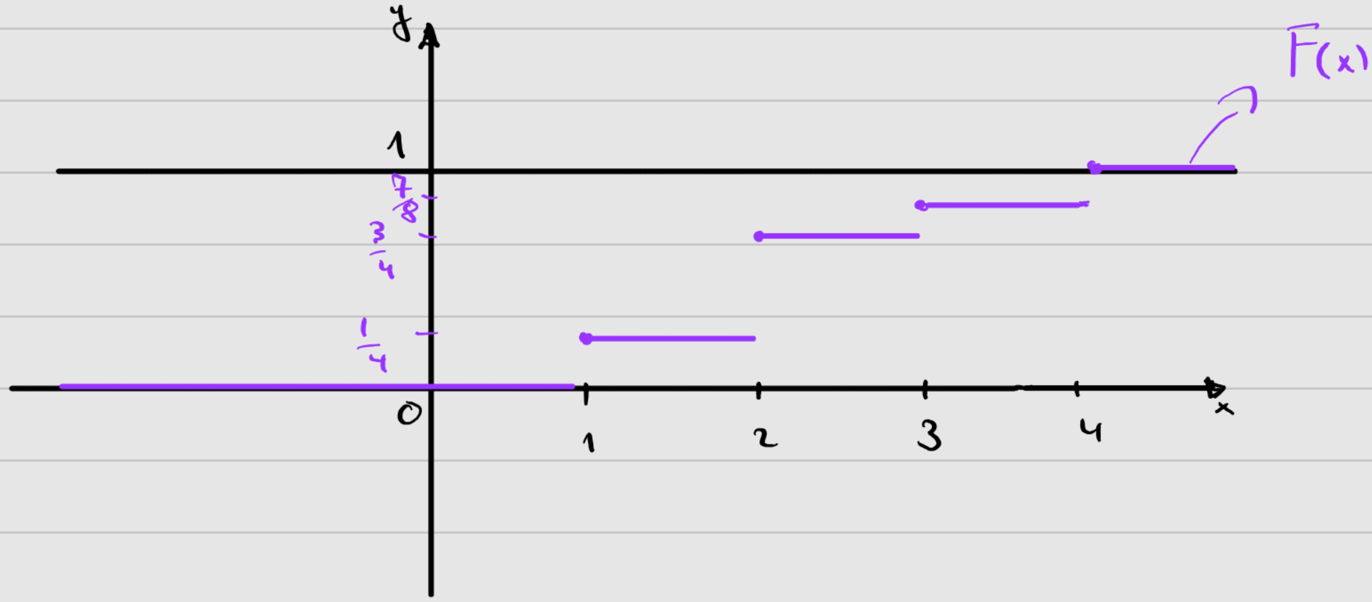
Sol: From the definitions we have:

$$F(r) = P(X \leq r) = \sum_{a \leq r} P(X=a) = \sum_{a \leq r} p(a),$$

hence: for $r < 1$: $F(r) = 0$, for $1 \leq r < 2$

$F(r) = p(1) = \frac{1}{4}$, for $2 \leq r < 3$ $F(r) = p(1) + p(2) = \frac{3}{4}$,

for $3 \leq r < 4$ $F(r) = p(1) + p(2) + p(3) = \frac{1}{4} + \frac{1}{2} + \frac{1}{8} = \frac{7}{8}$, and
for $4 \leq r$ $F(r) = p(1) + p(2) + p(3) + p(4) = 1$.
The graph is:



If X is a (discrete) random variable with probability mass function $p(x)$, then the **expectation**, or **expected value** of X is defined by:

$$E[X] = \sum_{x : p(x) > 0} x \cdot p(x).$$

P8 Find $E[X]$, where X is the random variable defined as the outcome when we roll a fair die.

Sol: For $i \in \{1, 2, \ldots, 6\}$, we have:

$$p(i) = P(X = i) = P(\text{roll the side with number } i) = \frac{1}{6}$$

since the die is fair. From this we deduce:

$$E[X] = 1 \cdot p(1) + 2 \cdot p(2) + 3 \cdot p(3) + 4 \cdot p(4) + 5 p(5) + 6 \cdot p(6)$$

$$= \frac{1}{6}(1+2+\ldots+6) = \frac{1}{6}\cdot 21 = \frac{7}{2} = 3.5 .$$

P9 A school class of 120 students is driven in 3 buses. There are 36 students in one bus, 40 in another, and 44 in the third bus. Later, one of the 120 students is randomly chosen. Let X denote the number of students on the bus of that student. Find $E[X]$!

Sol: Since the randomly chosen student is equally likely to be any of the 120 students, it follows that:

$$P(X=36) = P(\text{the student is from the first bus}) = \frac{36}{120},$$

and similarly $P(X=40) = \frac{40}{120}$ and $P(X=44) = \frac{44}{120}$.

So, the expected value of X is:

$$E[X] = \sum_x x\cdot p(x) = 36\cdot\frac{\overset{3}{\cancel{36}}}{\underset{12}{\cancel{120}}} + 40\cdot\frac{\overset{1}{\cancel{40}}}{\underset{3}{\cancel{120}}} + 44\cdot\frac{\overset{11}{\cancel{44}}}{\underset{30}{\cancel{120}}}$$

$$= \frac{1208}{30} = 40.2667.$$

P10 Let X denote a random variable that takes values -1, 0 and 1 with the following probabilities:

$$P(X=-1) = 0.2, \quad P(X=0) = 0.5 \text{ and } P(X=1) = 0.3 .$$

Compute $E[X]$ and $E[X^2]$.

Sol: Note that X^2 takes values 0 and 1 and $P(X^2=0) = P(X=0) = 0.5$, and $P(X^2=1) = P(X=1) + P(X=-1) = 0.3+0.2=0.5$, so

$E[X^2] = 0\cdot 0.5 + 1\cdot 0.5 = 0.5$. On the other hand:

$E[X] = (-1)\cdot 0.2 + 0\cdot 0.5 + 1\cdot 0.3 = 0.1$. (Notice that $E[X^2] \neq (E[X])^2$.)