Department of Computer Science, University of Verona
Degree in Applied Mathematics
Probability, Tuesday, July 19^{th}, 2016

Exercises

1. (a) A medical laboratory test ensures with a probability 0.95 the detection of a certain disease D, when the disease effectively exists. Meanwhile, the test also indicates a positive result for one percent of the people without the disease that have tried the test. We assume that 0.5 percent of the population has the disease D. We choose at random a person for which the test was positive. What is the probability that the person is ill ?

   (b) An aircraft engine fails with probability $1 - p$ where $p \in (0, 1)$, independently of the other engines. To complete its flight, the aircraft needs that the majority of its engines works. Calculate the probability to successfully complete the flight for an aircraft with 3 engines (i.e. at least 2 engines work).

   Solution:

   (a) We choose randomly a person. Let us denote the events

$$\begin{aligned} D &= \{\text{the person has the disease}\} \\ + &= \{\text{the test for this person is positive}\}. \end{aligned}$$

   Hence we know that

$$\begin{aligned} &\mathbb{P}(+ \mid D) = 0.95 \\ &\mathbb{P}\left(+ \mid \bar{D}\right) = 0.01 \\ &\mathbb{P}(D) = 0.5\% = 0.005 \end{aligned}$$

   We should compute the probability

$$\mathbb{P}(D \mid +).$$

   Using Bayes' formula we deduce

$$\begin{aligned} \mathbb{P}(D \mid +) &= \frac{\mathbb{P}(D \cap +)}{\mathbb{P}(+)} = \frac{\mathbb{P}(+ \mid D) \cdot \mathbb{P}(D)}{\mathbb{P}(+ \mid D) \cdot \mathbb{P}(D) + \mathbb{P}\left(+ \mid \bar{D}\right) \cdot \mathbb{P}\left(\bar{D}\right)} \\ &= \frac{0.95 \cdot 0.005}{0.95 \cdot 0.005 + 0.01 \cdot 0.995} = 0.323. \end{aligned}$$

   (b) Let us denote the events

$$E_i = \{\text{the engine } i \text{ works}\}, \ i = \overline{1, 3}.$$

   We should compute the probability of the event

$$\begin{aligned} X &= \{\text{the successfully flight for an aircraft with 3 engines}\} \\ &= \{\text{at least 2 from 3 engines work}\}. \end{aligned}$$

   Hence

$$\mathbb{P}(E_i) = p, \ i = \overline{1, 3}$$

and

$$X=\left(\bar{E}_1\cap E_2\cap E_3\right)\cup\left(E_1\cap\bar{E}_2\cap E_3\right)\cup\left(E_1\cap E_2\cap\bar{E}_3\right)\cup\left(E_1\cap E_2\cap E_3\right).$$

Therefore, since the events $\left\{\bar{E}_1\cap E_2\cap E_3\right\},\left\{E_1\cap\bar{E}_2\cap E_3\right\},\left\{E_1\cap E_2\cap\bar{E}_3\right\},\left\{E_1\cap E_2\cap E_3\right\}$ are disjoint and $E_i\ ,\ i=\overline{1,3}$, are independent (hence also $\bar{E}_1,E_2,E_3$ are independent and so on),

$$\begin{aligned}&\mathbb{P}\left(X\right)=\mathbb{P}\left(\left(\bar{E}_1\cap E_2\cap E_3\right)\cup\left(E_1\cap\bar{E}_2\cap E_3\right)\cup\left(E_1\cap E_2\cap\bar{E}_3\right)\cup\left(E_1\cap E_2\cap E_3\right)\right)\\&=\mathbb{P}\left(\bar{E}_1\cap E_2\cap E_3\right)+\mathbb{P}\left(E_1\cap\bar{E}_2\cap E_3\right)+\mathbb{P}\left(E_1\cap E_2\cap\bar{E}_3\right)+\mathbb{P}\left(E_1\cap E_2\cap E_3\right)\\&=\mathbb{P}\left(\bar{E}_1\right)\cdot\mathbb{P}\left(E_2\right)\cdot\mathbb{P}\left(E_3\right)+\mathbb{P}\left(E_1\right)\cdot\mathbb{P}\left(\bar{E}_2\right)\cdot\mathbb{P}\left(E_3\right)+\mathbb{P}\left(E_1\right)\cdot\mathbb{P}\left(E_2\right)\cdot\mathbb{P}\left(\bar{E}_3\right)\\&\quad+\mathbb{P}\left(E_1\right)\cdot\mathbb{P}\left(E_2\right)\cdot\mathbb{P}\left(E_3\right)=q\cdot p\cdot p+p\cdot q\cdot p+p\cdot p\cdot q+p\cdot p\cdot p=p^3+3p^2q.\end{aligned}$$

2. Let $n,m\in\mathbb{N}^*$ and $p\in(0,1)$. Let the independent random variables $X\sim$ *Binomial* (n,p), $Y\sim$ *Binomial* (m,p), and the independent random variables X_i given by $\mathbb{P}\left(X_i=-1\right)=\mathbb{P}\left(X_i=1\right)=1/2$, with $i\in\mathbb{N}^*$.

   (a) Find the mean and the variance of X.

   (b) Find the type of the distribution of the random variable $X+Y$. Deduce the variance of $X+Y$.

   (c) Find the type of the distribution of the random variable $\bar{X}_k:=\dfrac{k+\sum_{i=1}^k X_i}{2}$.

Hint: the notation $X\sim$ *Binomial*(n,p) means that X is Binomial distributed with parameters $n\in\mathbb{N}^*$ and $p\in(0,1)$, i.e. $\mathbb{P}(X=k)=C_n^kp^kq^{n-k}$, with $k=\overline{0,n}$, where $C_n^k=\frac{n!}{k!(n-k)!}$. It is useful also the expansion $(a+b)^n=\sum_{k=0}^n C_n^ka^kb^{n-k}$ and the relation $\sum_{i=0}^k C_n^iC_m^{k-i}=C_{n+m}^k$.

Solution:

(a) Since

$$X:\begin{pmatrix}k\\C_n^kp^kq^{n-k}\end{pmatrix}_{k=\overline{0,n}},$$

we have

$$\begin{aligned}&\mathbb{E}\left(X\right)=\sum\nolimits_{k=0}^n k\cdot C_n^kp^kq^{n-k}=\sum\nolimits_{k=1}^n k\cdot C_n^kp^kq^{n-k}=\sum\nolimits_{k=1}^n\frac{n!}{(k-1)!\,(n-k)!}p^kq^{n-k}\\&=np\sum\nolimits_{k=1}^n\frac{(n-1)!}{(k-1)!\,[(n-1)-(k-1)]!}\,p^{k-1}q^{(n-1)-(k-1)}\\&=np\sum\nolimits_{k'=0}^{n-1}\frac{(n-1)!}{k'!\,[(n-1)-k']!}\,p^{k'}q^{(n-1)-k'}=np\left(p+q\right)^{n-1}=np\end{aligned}$$

(we denoted $k'=k-1$ and we used the Newton expansion $(p+q)^{n-1}=\sum_{k'=0}^{n-1}\frac{(n-1)!}{k'![(n-1)-k']!}\,p^{k'}q^{(n-1)-k'}$).

Also

$$\mathbb{E}\left(X^2\right) = \sum\nolimits_{k=0}^{n} k^2 \cdot C_n^k p^k q^{n-k} = \sum\nolimits_{k=1}^{n} k^2 \cdot C_n^k p^k q^{n-k} = \sum\nolimits_{k=1}^{n} k \, \frac{n!}{(k-1)!\,(n-k)!} p^k q^{n-k}$$

$$= \sum\nolimits_{k=1}^{n} (k-1+1) \, \frac{n!}{(k-1)!\,(n-k)!} p^k q^{n-k}$$

$$= \sum\nolimits_{k=1}^{n} (k-1) \, \frac{n!}{(k-1)!\,(n-k)!} p^k q^{n-k} + \sum\nolimits_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k q^{n-k}$$

$$= \sum\nolimits_{k=2}^{n} \frac{n!}{(k-2)!\,(n-k)!} p^k q^{n-k} + \sum\nolimits_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k q^{n-k}$$

$$= n\,(n-1)\,p^2 \sum\nolimits_{k=2}^{n} \frac{(n-2)!}{(k-2)!\,[(n-2)-(k-2)]!} p^{k-2} q^{(n-2)-(k-2)}$$

$$+ np \sum\nolimits_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,[(n-1)-(k-1)]!} p^{k-1} q^{(n-2)-(k-2)}$$

$$= n\,(n-1)\,p^2 \sum\nolimits_{k'=0}^{n-2} \frac{(n-2)!}{k'!\,[(n-2)-k']!} p^{k'} q^{(n-2)-k'} + np \sum\nolimits_{k''=0}^{n-1} \frac{(n-1)!}{k''!\,[(n-1)-k'']!} p^{k''} q^{(n-2)-k''}$$

$$= n\,(n-1)\,p^2\,(p+q)^{n-2} + np\,(p+q)^{n-1} = n\,(n-1)\,p^2 + np.$$

(we denoted $k' = k-1$, $k'' = k-2$ and we used the Newton expansion).

Hence

$$\operatorname{Var}(X) = \mathbb{E}\left(X^2\right) - [\mathbb{E}(X)]^2 = n\,(n-1)\,p^2 + np - (np)^2 = np - np^2 = np\,(1-p) = npq.$$

(b) If $X \sim$ *Binomial* (n,p) and $Y \sim$ *Binomial* (m,p), then $X(\Omega) = \{0,1,2,\ldots,n\}$ and $Y(\Omega) = \{0,1,2,\ldots,m\}$ and therefore all the values of the r.v. $(X+Y)$ are

$$(X+Y)(\Omega) = \{0,1,2,\ldots,n+m\}.$$

We compute now

$$\mathbb{P}(X+Y=k) = \mathbb{P}\left(\bigcup\nolimits_{i=0}^{k} \{X=i, Y=k-i\}\right) = \sum\nolimits_{i=0}^{k} \mathbb{P}(X=i, Y=k-i)$$

$$= \sum\nolimits_{i=0}^{k} \mathbb{P}(\{X=i\} \cap \{Y=k-i\}) = \sum\nolimits_{i=0}^{k} \mathbb{P}(X=i) \cdot \mathbb{P}(Y=k-i)$$

$$= \sum\nolimits_{i=0}^{k} C_n^i p^i q^{n-i} \cdot C_m^{k-i} p^{k-i} q^{n-k+i} = p^k q^{n-k} \sum\nolimits_{i=0}^{k} C_n^i \cdot C_m^{k-i} = C_{n+m}^k \cdot p^k q^{n-k},$$

since we have the algebraic identity

$$\sum\nolimits_{i=0}^{k} C_n^i C_m^{k-i} = C_{n+m}^k.$$

Hence, we obtain that

$$(X+Y) \sim \textit{Binomial}\,(n+m,p).$$

(c) We see that

$$1+X_i : \begin{pmatrix} 0 & 2 \\ 1/2 & 1/2 \end{pmatrix} \quad \Rightarrow \quad \frac{1+X_i}{2} : \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}$$

hence

$$\frac{1+X_i}{2} \sim \textit{Binomial}\,(1,1/2) = \textit{Bernoulli}\,(1/2).$$

We have

$$\bar{X}_k = \frac{k + \sum_{i=1}^{k} X_i}{2} = \frac{\sum_{i=1}^{k} (1+X_i)}{2} = \sum\nolimits_{i=1}^{k} \frac{1+X_i}{2} \sim \textit{Binomial}\,(k,1/2)$$

(we should use point (b) in order to make the sum of k binomial distributions).

3. Let the function $F : \mathbb{R} \to \mathbb{R}$ be given by

$$F(x) = \begin{cases} 1 + a, & \text{if } x < 0, \\ b, & \text{if } x \in [0, 2), \\ c + 2b & \text{if } x \in [2, 3), \\ d & \text{if } x \geq 3. \end{cases}$$

(a) Under which conditions on a, b, c, d function F is a cumulative distribution function?

(b) If F is a cumulative distribution function associated to a random variable X, then compute $\mathbb{P}(X > 2.5)$ and $\mathbb{P}(-1 < X \leq 1.5)$, using F.

(c) Find a, b, c, d if we know that $\mathbb{P}(X > 2.5) = 0.25$ and $\mathbb{P}(-1 < X \leq 1.5) = 0.25$.

(d) Find the law (the table) of X.

Solution:

(a) F is a cumulative distribution function if

$$\begin{aligned} (i) &\quad F \text{ is increasing} \\ (ii) &\quad F \text{ is right continuous} \\ (iii) &\quad \lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to +\infty} F(x) = 1. \end{aligned}$$

Hence we should impose

$$a = -1, \quad d = 1$$

and

$$0 \leq b \leq c + 2b \leq 1 \quad \Rightarrow \quad 0 \leq b \leq c + 2b \leq 1.$$

(b) Using the definition

$$F_X(a) = \mathbb{P}(X \leq a)$$

we obtain some standard formulas

$$\begin{aligned} \mathbb{P}(X > a) &= 1 - F_X(a), \\ \mathbb{P}(a < X \leq b) &= F_X(b) - F_X(a), \\ \mathbb{P}(X = a) &= F_X(a) - F_X(a - 0), \\ \mathbb{P}(a \leq X < b) &= F_X(b - 0) - F_X(a - 0). \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{P}(X > 2.5) &= 1 - F_X(2.5) = 1 - (c + 2b), \\ \mathbb{P}(-1 < X \leq 1.5) &= F_X(1.5) - F_X(-1) = b - (1 + a) = b. \end{aligned}$$

(c) Since $\mathbb{P}(X > 2.5) = 0.25$ and $\mathbb{P}(-1 < X \leq 1.5) = 0.25$, we have

$$1 - (c + 2b) = 0.25, \quad b = 0.25.$$

Hence

$$b = 0.25, \quad c = 0.25.$$

(d) We obtain that

$$\mathbb{P}(X = a) = F_X(a) - F_X(a - 0) = 0$$

for any $a \in \mathbb{R}$ a discontinuous point for F. Therefore $\mathbb{P}(X = a) = 0$, for any $a \notin \{0, 2, 3\}$.

After that

$$\mathbb{P}(X=0) = F_X(0) - F_X(0-0) = b - (1+a) = 0.25,$$
$$\mathbb{P}(X=2) = F_X(2) - F_X(2-0) = c + 2b - b = 0.5,$$
$$\mathbb{P}(X=3) = F_X(3) - F_X(3-0) = d - (c+2b) = 0.25.$$

Therefore

$$X : \begin{pmatrix} 0 & 2 & 3 \\ 0.25 & 0.5 & 0.25 \end{pmatrix}.$$