II Università di Roma, Tor Vergata
Dipartimento d'Ingegneria Civile e Ingegneria Informatica
LM in Ingegneria dell'Informazione e dell'Automazione
Complementi di Probabilità e Statistica - Advanced Statistics
Instructors: Roberto Monte & Massimo Regoli
Solved Problems on Random Variables 2022-12-08

Problem 1 *Let X be a geometrically distributed random variable with success probability p.*

1. *Are the moments of order 1 and 2 of X finite?*
2. *If the moments of order 1 and 2 are finite, can you compute $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$.*
3. *Assume that $p = 0.1$. Cand you apply the Tchebychev inequality to estimate $\mathbf{P}(-2 < X < 23)$.*

Solution. . $\square$

Problem 2 *Let X be a uniform continuous real random variable with states in the interval $(0,1)$. In symbols, $X \sim Unif(0,1)$. Let $a, b \in \mathbb{R}$ such that $a < b$. Prove that the random variable $Y \overset{def}{=} a + (b-a)X$ is a uniform continuous real random variable with states in the interval (a,b). In symbols, $Y \sim Unif(a,b)$.*

Solution. .

Problem 3 *Let $X_1, \ldots, X_n$ (totally) independent random variables uniformly distributed in the interval $(0,1)$. Hence, $X_1, \ldots, X_n$ are absolutely continuous with density $f_{X_k} : \mathbb{R} \to \mathbb{R}_+$ given by*

$$f_{X_k}(x) \overset{def}{=} 1_{(0,1)}(x) \quad \forall x \in \mathbb{R},$$

on varying of $k = 1, \ldots, n$. Consider the random variables

$$\check{X}_n \overset{def}{=} \max(X_1, \ldots, X_n) \quad \text{and} \quad \hat{X}_n \overset{def}{=} \min(X_1, \ldots, X_n).$$

Compute $\mathbf{E}[\check{X}_n]$ and $\mathbf{E}\left[\hat{X}_n\right]$.

Exercise 4 *Hint: it might be useful to determine the distribution functions $F_{\check{X}_n} : \mathbb{R} \to \mathbb{R}_+$ and $F_{\hat{X}_n} : \mathbb{R} \to \mathbb{R}_+$.*

Solution. .

Problem 5 *The decoration of a Christmas tree in a mall is made by 1.000 small light bulbs. The lifetime of each light bulb is exponentially distributed with an average life time of 20 days (rather cheap bulbs indeed!). The mall manager decides to turn on the lights of the Christmas tree on the midnight of the 15-th of November. Estimate the probability that at least 800 bulbs are still working on the midnight of the 25-th of December.*

Hints: compute the probability that each bulb will last until the midnight of the 25-th of December; write X for the random variable counting the number of light bulbs out of 1.000 which are still on at the midnight of the 25-th of December and guess how it is distributed; use the Markov inequality to make the estimate.

Solution. Let $T : \Omega \to \mathbb{R}_+$ be the random variable representing the life time of each bulb. Under our assumptions T is exponentially distributed with rate λ. That is T is absolutely continuous with density

$$f_T(x) \overset{\text{def}}{=} \lambda e^{-\lambda x} 1_{\mathbb{R}_+}(x), \quad \forall x \in \mathbb{R}.$$

Recall that

$$\begin{aligned}\int_{\mathbb{R}} f_T(x)\, d\mu_L(x) &= \int_{\mathbb{R}} \lambda e^{-\lambda x} 1_{\mathbb{R}_+}(x)\, d\mu_L(x) = \int_{\mathbb{R}_+} \lambda e^{-\lambda x} d\mu_L(x) = \int_0^{+\infty} \lambda e^{-\lambda x} dx \\ &= \lim_{x\to+\infty} \int_0^x \lambda e^{-\lambda u} du = -\lim_{x\to+\infty} \int_0^x e^{-\lambda u} d(-\lambda u) = -\lim_{x\to+\infty} \int_0^{-\lambda x} e^v dv \\ &= \lim_{x\to+\infty} e^v\big|_{-\lambda x}^0 = \lim_{x\to+\infty} \left(1 - e^{-\lambda x}\right) = 1.\end{aligned}$$

Recall also that

$$\mathbf{E}[X] = \int_{\mathbb{R}} x f_T(x)\, d\mu_L(x) = \int_{\mathbb{R}_+} \lambda x e^{-\lambda x} d\mu_L(x) = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = \lim_{x\to+\infty} \int_0^x \lambda u e^{-\lambda u} du.$$

Now, integrating by parts,

$$\begin{aligned}\int_0^x \lambda u e^{-\lambda u} du &= -\int_0^x u\, d\left(e^{-\lambda u}\right) = -\ u e^{-\lambda u}\Big|_0^x + \int_0^x e^{-\lambda u} du = -\ u e^{-\lambda u}\Big|_0^x - \frac{1}{\lambda}\int_0^x e^{-\lambda u} d(-\lambda u) \\ &= -\ u e^{-\lambda u}\Big|_0^x - \frac{1}{\lambda}\int_0^{-\lambda x} e^v dv = -\ u e^{-\lambda u}\Big|_0^x + \frac{1}{\lambda}\, e^v\big|_{-\lambda x}^0 \\ &= -x e^{-\lambda x} + \frac{1}{\lambda}\left(1 - e^{-\lambda x}\right)\end{aligned}$$

It follows,

$$\mathbf{E}[X] = \frac{1}{\lambda}.$$

Hence, since the average life time of each bulb is 20 days, we set

$$\frac{1}{\lambda} = 20 \Leftrightarrow \lambda = \frac{1}{20}.$$

Then, we then have

$$\begin{aligned}\mathbf{P}(T \le t) &= \int_{(-\infty,t]} f_T(x)\, d\mu_L(x) = \int_{(-\infty,t]} \frac{1}{20} e^{-\frac{1}{20}x} 1_{\mathbb{R}_+}(x)\, d\mu_L(x) \\ &= \int_{(0,t]} \frac{1}{20} e^{-\frac{1}{20}x} d\mu_L(x) = \int_0^t \frac{1}{20} e^{-\frac{1}{20}x} dx = -\int_0^t e^{-\frac{1}{20}x} d\left(-\frac{1}{20}x\right) \\ &= -\int_0^{-\frac{t}{20}} e^u du = e^u\big|_{-\frac{t}{20}}^0 = 1 - e^{-\frac{t}{20}},\end{aligned}$$

for any $t \ge 0$. We are interested to compute the probability that a bulb is still working after 40 days it has been turned on. Hence, we have to set $t = 40$ and compute

$$\mathbf{P}(T > 40) = 1 - \mathbf{P}(T \le 40) = 1 - \left(1 - e^{-2}\right) = e^{-2} \simeq 0.13534.$$

Now, let $X : \Omega \to \mathbb{N}_0$ be the random variable representing the number of bulbs which are still working after 40 days. Since the bulbs in the circuit are parallel connected, we can assume that they are independent from each other. Therefore, X is a binomial random variable with parameters $n = 1000$ and $p = e^{-2}$. As a consequence, thanks to the Markov inequality, we can write

$$\mathbf{P}(X \geq 800) \leq \frac{\mathbf{E}[X]}{800} = \frac{1000e^{-2}}{800} = \frac{5}{4}e^{-2} \simeq 0.16917.$$

However, we have also

$$\mathbf{P}(X \geq 800) = 1-\mathbf{P}(X < 800) = 1-\sum_{k=0}^{799}\mathbf{P}(X = k) = 1-\sum_{k=0}^{799}\binom{1000}{k}e^{-2k}\left(1-e^{-2}\right)^{1000-k} \simeq 2.64233\cdot 10^{-14},$$

which shows that the bound obtained via the Markov inequality may be rather loose. □

Problem 6 *Let X be a geometrically distributed random variable with success probability $p = 0.1$. Compute $\mathbf{E}[X]$ and $\mathbf{D}^2[X]$. Use the Tchebychev inequality to estimate $\mathbf{P}(-2 < X < 23)$.*

Solution. . □

Exercise 7 *An empathic professor aims to help his students to pass the hard final exam of his course in Probability and Statistics. To this goal, he splits the course program in two parts and gives his students an intermediate written test on the first part of the course. Assume that*

- *the 18% of the students attending the course who pass the final exam on their first try got a mark not lower than 25 in the intermediate test;*
- *the 24% of the students attending the course who pass the final exam on their first try got a mark in the range $20 - 24$ in the intermediate test;*
- *the 30% of the students attending the course who pass the final exam on their first try got a mark not higher than 19 in the intermediate test;*
- *the 4% of the students attending the course who pass the final exam on their first try did not take the intermediate test.*

Assume also that

- *the 20% of the students attending the course get a mark not lower than 25 in the intermediate test;*
- *the 30% of the students attending the course get a mark in the range $20 - 24$ in the intermediate test;*
- *the 10% of the students attending the course do not take the intermediate test.*

Compute:

1. *the probability that a student attending the course passes the final exam on her first try, given that she got a mark not lower than 25 in the intermediate test;*
2. *the probability that a student attending the course passes the final exam on her first try, given that she got a mark not higer than 19 in the intermediate test;*
3. *the probability that a student attending the course passes the final exam on her first try, given that she did not take the intermediate test;*

4. *the probability that a student attending the course passes the final exam on her first try;*
5. *the probability that a student attending the course got a mark not lower than* 25 *in the intermediate test, given that she passes the final exam on her first try;*
6. *the probability that a student attending the course got a mark not lower than* 25 *in the intermediate test, given that she does not pass the final exam on her first try;*
7. *the probability that a student attending the course does not pass the final exam on her first try given that she did not take the intermediate test.*

   Hint: write $T_{\geq 25}$ [resp. T_{20-24}, $T_{\leq 19}$] for the event "a randomly choosen student attending the course gets a mark not lower than 25 [resp. in the range $20 - 24$, not higher than 19] in the intermediate test". Write also T_0 for the event "a randomly choosen student attendin

Solution. .

Problem 8 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $X, Y \in \mathcal{L}^2(\Omega; \mathbb{R})$. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}_+$ given by*

$$f(a,b) \stackrel{def}{=} \mathbf{E}\left[(X - (a + bY))^2\right].$$

Prove that there exists

$$(a^*, b^*) \equiv \underset{(a,b)\in\mathbb{R}^2}{\arg\min}\{f(a,b)\}$$

and compute it.

Solution. By virtue of the properties of the expectation operator, a straightforward computation yields

$$f(a,b) = a^2 + 2ab\mathbf{E}[Y] + b^2\mathbf{E}\left[Y^2\right] - 2a\mathbf{E}[X] - 2b\mathbf{E}[XY] + \mathbf{E}\left[X^2\right].$$

Therefore, $f(a,b)$ is a second order polynomial in the variables a and b. Consider the zeroes of the partial derivatives of $f(a,b)$, we have

$$\begin{aligned}
\partial f_a(a,b) = 0 &\Leftrightarrow a + b\mathbf{E}[Y] = \mathbf{E}[X], \\
\partial f_b(a,b) = 0 &\Leftrightarrow a\mathbf{E}[Y] + b\mathbf{E}\left[Y^2\right] = \mathbf{E}[XY].
\end{aligned}$$

Thus, a point (a,b) is a candidate local minimum only if

$$\begin{aligned}
a &= \frac{\begin{vmatrix} \mathbf{E}[X] & \mathbf{E}[Y] \\ \mathbf{E}[XY] & \mathbf{E}\left[Y^2\right] \end{vmatrix}}{\begin{vmatrix} 1 & \mathbf{E}[Y] \\ \mathbf{E}[Y] & \mathbf{E}\left[Y^2\right] \end{vmatrix}} = \frac{\mathbf{E}[X]\,\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]\,\mathbf{E}[XY]}{\mathbf{D}^2[Y]} \\
&= \frac{\mathbf{E}[X]\,\mathbf{E}\left[Y^2\right] - \mathbf{E}[X]\,\mathbf{E}[Y]^2 + \mathbf{E}[X]\,\mathbf{E}[Y]^2 - \mathbf{E}[Y]\,\mathbf{E}[XY]}{\mathbf{D}^2[Y]} \\
&= \frac{\mathbf{E}[X]\left(\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2\right) + \mathbf{E}[Y]\left(\mathbf{E}[X]\,\mathbf{E}[Y] - \mathbf{E}[XY]\right)}{\mathbf{D}^2[Y]} \\
&= \mathbf{E}[X] - \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\mathbf{E}[Y], \\
b &= \frac{\begin{vmatrix} 1 & \mathbf{E}[X] \\ \mathbf{E}[Y] & \mathbf{E}[XY] \end{vmatrix}}{\begin{vmatrix} 1 & \mathbf{E}[Y] \\ \mathbf{E}[Y] & \mathbf{E}\left[Y^2\right] \end{vmatrix}} = \frac{\mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y]}{\mathbf{D}^2[Y]} = \frac{cov(X,Y)}{\mathbf{D}^2[Y]}.
\end{aligned}$$

Moreover, we have

$$\partial^2 f_{a,a}(a,b) = 2 > 0, \qquad \partial^2 f_{a,b}(a,b) = 2\mathbf{E}[Y], \qquad \partial^2 f_{b,b}(a,b) = 2\mathbf{E}\left[Y^2\right] > 0.$$

Hence, the Hessian matrix Hf at (a,b) is given by

$$(Hf)(a,b) = \begin{pmatrix} 2 & 2\mathbf{E}[Y] \\ 2\mathbf{E}[Y] & 2\mathbf{E}\left[Y^2\right] \end{pmatrix}$$

and has determinant

$$\det((Hf)(a,b)) = 4\left(\mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]\right) = 4\mathbf{D}^2[Y] > 0.$$

It follows that the point

$$(a^*, b^*) = \left(\mathbf{E}[X] - \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\mathbf{E}[Y], \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\right)$$

is actually a local minimum. On the other hand, it is not difficult to show that $f(a,b)$ is a convex function. In fact, since

$$0 < \mathbf{D}^2[Y] = \mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2,$$

we have

$$4\mathbf{E}[Y]^2 \leq 4\mathbf{E}\left[Y^2\right]$$

and

$$\left(\mathbf{E}\left[Y^2\right] + 1\right)^2 \geq 4\mathbf{E}[Y]^2 + \left(\mathbf{E}\left[Y^2\right] - 1\right)^2.$$

The latter implies

$$0 \leq \mathbf{E}\left[Y^2\right] + 1 - \sqrt{4\mathbf{E}[Y]^2 + (\mathbf{E}[Y^2] - 1)^2} \leq \mathbf{E}\left[Y^2\right] + 1 + \sqrt{4\mathbf{E}[Y]^2 + (\mathbf{E}[Y^2] - 1)^2},$$

that is to say, the eigenvalues the Hessian matrix $(Hf)(a,b)$ are positive. As a consequence,

$$(a^*, b^*) = \underset{(a,b)\in\mathbb{R}^2}{\arg\min}\{f(a,b)\}.$$

Note that the first order Y-polynomial

$$\phi(Y) \overset{\text{def}}{=} \mathbf{E}[X] - \frac{cov(X,Y)}{\mathbf{D}^2[Y]}\mathbf{E}[Y] + \frac{cov(X,Y)}{\mathbf{D}^2[Y]}Y$$

turns out to be the best $(\sigma(Y), \mathcal{B}(\mathbb{R}))$-random first order Y-polynomial which approximates the $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$-random variable X w.r.t. the square norm.

Problem 9 *Students are allowed to take a test twice in an examination session. Assume that* 8 *students over* 10 *pass the test the first time. For those who fail or do not show up, only* 6 *students over* 10 *pass the test the second time.*

1. *Find the probability that a randomly selected student (who needs to pass the test) passes the test.*
2. *Assuming that a student passed the test what is the probability she passed on the first try?*
3. *Consider that a part of the first test presented the following problem: two dice are rolled and the number of the upper faces are observed. Is the event "the sum of the observed numbers is* 7*" independent of the event "the number observed on the upper face of a die is* 5*"? Could you give a solution to this problem?*

Solution. . $\square$

Problem 10 *There is a group of n persons who checked their hat at a theatre. When they went to take their hats back the hatter went mad and started giving random persons random hats. What is the expected amount of persons who get their hat back? What is the probability that everyone gets their hat back?*

Hint: consider the k-th person of the group, for $k = 1, \ldots, n$, and write C_k for the random variable expressing the circumstance that the k-th person gets her hat back.

Solution. . $\square$

Problem 11 (Sheldon M. Ross - 4.2 - 4.3) *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let $(X_n)_{n\geq 1}$ be a sequence of independent Bernoulli random variables. Recall that for a Bernoulli random variable X we have*

$$X = \begin{cases} 1 & \mathbf{P}(X = 1) = p \\ 0 & \mathbf{P}(X = 0) = q \end{cases},$$

where $p \in (0, 1)$ and $q \equiv 1 - p$. Consider the sequence $(Z_n)_{n\geq 1}$ of random variables given by

$$Z_n \overset{def}{=} \textstyle\sum_{k=1}^{n} X_k, \quad \forall n \geq 1$$

and let $(H_n)_{n\geq 1}$ be the sequence of random variables given by

$$H_n \overset{def}{=} 2Z_n - n, \quad \forall n \geq 1.$$

1. *Assume that X_n is the random variable which represents the toss of a coin with the convention that "success" [resp. "failure"] is for the outcome "heads" [resp. "tails"] represented, in turn, by the outcome 1 [resp. 0]. Give an interpretation of the random variables Z_n and H_n and compute their mean and variance.*

2. *Assume that $p = 1/2$. Prove that we have*

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{H_n}{\sqrt{n}} < z\right) = \Phi(z),$$

for every $z \in \mathbb{R}$, where Φ is the standard notation for the distribution function of the standard normal random variable.

Solution.

1. By virtue of the meaning of the random variable X_n it is clearly seen that the random variable Z_n represents the number of successes $k = 0, 1, \ldots, n$ that we may obtain in n tosses and

$$H_n \overset{\text{def}}{=} 2Z_n - n = Z_n - (n - Z_n)$$

represents the number of heads minus the number of tails that we obtain in n tosses, namely the number of successes minus the number of failures that we may obtain in n tosses. Note that Z_n is binomially distributed with number of trials parameter n and success parameter p. That is

$$\mathbf{P}(Z_n = k) = \binom{n}{k} p^k q^{n-k}$$

for every $k = 0, 1, \ldots, n$. Now, we have

$$\mathbf{E}[X_n] = 1 \cdot \mathbf{P}(X = 1) + 0 \cdot \mathbf{P}(X = 0) = p$$

and

$$\mathbf{E}\left[X_n^2\right] = 1^2 \cdot \mathbf{P}(X = 1) + 0^2 \cdot \mathbf{P}(X = 0) = p.$$

These imply

$$\mathbf{D}^2[X_n] = \mathbf{E}\left[X_n^2\right] - \mathbf{E}[X_n]^2 = p - p^2 = p(1-p) = pq.$$

As a consequence,

$$\mathbf{E}[Z_n] = \mathbf{E}\left[\textstyle\sum_{k=1}^n X_k\right] = \textstyle\sum_{k=1}^n \mathbf{E}[X_k] = \textstyle\sum_{k=1}^n p = np.$$

Furthermore, thanks to the independence of the random variables of the sequence $(X_n)_{n\geq 1}$,

$$\mathbf{D}^2[Z_n] = \mathbf{D}^2\left[\textstyle\sum_{k=1}^n X_k\right] = \textstyle\sum_{k=1}^n \mathbf{D}^2[X_k] = \textstyle\sum_{k=1}^n pq = npq.$$

, since sum of independent random variables which are Bernoulli distributed with success parameter p We then have

$$\mathbf{E}[H_n] = \mathbf{E}[2Z_n - n] = 2\mathbf{E}[Z_n] - n = 2np - n = n(2p-1)$$

and

$$\mathbf{D}^2[H_n] = \mathbf{D}^2[2Z_n - n] = 4\mathbf{D}^2[Z_n] = 4npq.$$

2. Under the additional assumption $p = 1/2$, we have

$$\mathbf{E}[Z_n] = \frac{1}{2}n, \quad \mathbf{D}^2[Z_n] = \frac{1}{4}n, \quad \mathbf{E}[H_n] = 0, \quad \mathbf{D}^2[H_n] = n.$$

It follows

$$\mathbf{P}\left(\frac{H_n}{\sqrt{n}} < z\right) = \mathbf{P}\left(\frac{2Z_n - n}{\sqrt{n}} < z\right) = \mathbf{P}\left(\frac{Z_n - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} < z\right) = \mathbf{P}\left(\frac{Z_n - \mathbf{E}[Z_n]}{\mathbf{D}[Z_n]} < z\right) \tag{1}$$

On the other hand, by the Central limit Theorem, we have

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{Z_n - \mathbf{E}[Z_n]}{\mathbf{D}[Z_n]} < z\right) = \Phi(z), \tag{2}$$

for every $z \in \mathbb{R}$. Combining (1) and (2), the desired claim immediately follows.

Problem 12 *Suppose that we roll a standard fair die* 100 *times. Let X be the sum of the numbers that appear over the* 100 *rolls. Use the Tchebychev inequality to bound* $\mathbf{P}[(|X - 350| \geq 50)$.

Solution. . □

Problem 13 *Let $(\Omega, \mathcal{E}, \mathbf{P}) \equiv \Omega$ be a probability space and let X and Y discrete real random variables on Ω. Write $X(\Omega) \equiv \{x_m\}_{m\in M}$ and $Y(\Omega) \equiv \{y_n\}_{n\in N}$, where $M, N \subseteq \mathbb{N}$. Prove that X and Y are independent if and only if*

$$\mathbf{P}(X = x_m, Y = y_n) = \mathbf{P}(X = x_m)\mathbf{P}(Y = y_n), \quad \forall (m,n) \in M \times N. \tag{3}$$

Solution. The condition is clearly necessary. To prove that it is also sufficient, consider any couple $B, C \in \mathcal{B}(\mathbb{R})$. We have

$$\{X \in B\} = \bigcup_{m|x_m \in B} \{X = x_m\} \quad \text{and} \quad \{Y \in C\} = \bigcup_{n|y_n \in C} \{Y = y_n\}.$$

Therefore,

$$\begin{aligned}
\{X \in B, Y \in C\} &= \{X \in B\} \cap \{Y \in C\} \\
&= \left(\bigcup_{m|x_m \in B} \{X = x_m\} \right) \cap \left(\bigcup_{n|y_n \in C} \{Y = y_n\} \right) \\
&= \bigcup_{m|x_m \in B,\ n|y_n \in C} \{X = x_m\} \cap \{Y = y_n\} \\
&\equiv \bigcup_{m|x_m \in B,\ n|y_n \in C} \{X = x_m, Y = y_n\},
\end{aligned}$$

where the events of the family $\{\{X = x_m, Y = y_n\}\}_{m|x_m \in B,\ n|y_n \in C}$ are pairwise incompatible. As a consequence, under Assumption (3), we have

$$\begin{aligned}
\mathbf{P}(X \in B, Y \in C) &= \mathbf{P}\left(\bigcup_{m|x_m \in B,\ n|y_n \in C} \{X = x_m, Y = y_n\} \right) \\
&= \sum_{m|x_m \in B\ ,n|y_n \in C} \mathbf{P}(X = x_m, Y = y_n) \\
&= \sum_{m|x_m \in B\ ,n|y_n \in C} \mathbf{P}(X = x_m)\, \mathbf{P}(Y = y_n) \\
&= \left(\sum_{m|x_m \in B} \mathbf{P}(X = x_m) \right) \left(\sum_{n|y_n \in C} \mathbf{P}(Y = y_n) \right) \\
&= \mathbf{P}\left(\bigcup_{m|x_m \in B} \{X = x_m\} \right) \mathbf{P}\left(\bigcup_{n|y_n \in C} \{Y = y_n\} \right) \\
&= \mathbf{P}(X \in B)\, \mathbf{P}(Y \in C).
\end{aligned}$$

This, thanks to the arbitrariness of $B, C \in \mathcal{B}(\mathbb{R})$, yields the independence of X and Y.

.