Probability Theory
2025/26 Block 2b
Instructor: Gilles Bonnet
Exam
Wednesday 24 June 2026
Duration: 2 hours

Solutions

1. A submitted file is produced by one of three independent software systems A, B, C. A randomly chosen file comes from these systems with probabilities

$$\frac{1}{2}, \qquad \frac{1}{3}, \qquad \frac{1}{6},$$

respectively. The probability that the file is rejected by the submission portal is

$$\frac{1}{20} \text{ for system } A, \qquad \frac{1}{10} \text{ for system } B, \qquad \frac{1}{5} \text{ for system } C.$$

Different files are produced independently of one another.

Give exact answers.

(a) (4 points) What is the probability that a randomly chosen file is rejected?

Solution: Let R be the event that the file is rejected. By the law of total probability,

$$\begin{aligned} \mathbb{P}(R) &= \mathbb{P}(R \mid A)\mathbb{P}(A) + \mathbb{P}(R \mid B)\mathbb{P}(B) + \mathbb{P}(R \mid C)\mathbb{P}(C) \\ &= \frac{1}{2}\frac{1}{20} + \frac{1}{3}\frac{1}{10} + \frac{1}{6}\frac{1}{5} \\ &= \frac{1}{40} + \frac{1}{30} + \frac{1}{30} = \frac{11}{120}. \end{aligned}$$

(b) (4 points) Given that a file is rejected, what is the probability that it came from system C?

Solution: By Bayes' formula,

$$\mathbb{P}(C \mid R) = \frac{\mathbb{P}(R \mid C)\mathbb{P}(C)}{\mathbb{P}(R)} = \frac{\frac{1}{5}\frac{1}{6}}{\frac{11}{120}} = \frac{4}{11}.$$

(c) (4 points) Five files are submitted. What is the probability that exactly two of them are rejected?

Solution: For five independent files, the number rejected has binomial distribution with parameters 5 and 11/120. If $Z \sim \mathrm{Bin}(n, p)$, then

$$\mathbb{P}(Z = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Hence

$$\mathbb{P}(\text{exactly two rejected}) = \binom{5}{2} \left(\frac{11}{120}\right)^2 \left(\frac{109}{120}\right)^3.$$

(d) (4 points) A student repeatedly submits independently produced files until one is accepted. What is the expected number of submissions?

Solution: The acceptance probability of one file is $1 - 11/120 = 109/120$. The number of submissions until the first accepted file is geometric with success parameter 109/120. If $G \sim \mathrm{Geo}(p)$, then $\mathbb{E}(G) = 1/p$, so the expectation is

$$\frac{1}{109/120} = \frac{120}{109}.$$

2. Let T be a continuous random variable with density

$$f_T(t) = \begin{cases} c\,t(3-t), & 0 < t < 3, \\ 0, & \text{otherwise}, \end{cases}$$

where c is the normalizing constant. Let

$$Y = T^2.$$

Compute $\mathbb{E}(Y)$ and $\mathbb{V}\mathrm{ar}(Y)$. Give exact answers. Your computation should include the value of c.

(a) (4 points) Find the value of c.

Solution: The normalizing constant is determined by

$$\int_{-\infty}^{\infty} f_T(t)\,dt = 1.$$

Here this gives

$$1 = c\int_0^3 t(3-t)\,dt = c\frac{9}{2},$$

so $c = 2/9$.

(b) (5 points) Compute $\mathbb{E}(Y)$.

Solution: For a function g, the expectation formula is

$$\mathbb{E}(g(T)) = \int_{-\infty}^{\infty} g(t) f_T(t)\,dt.$$

Using $g(t) = t^2$, $c = 2/9$, and $Y = T^2$,

$$\mathbb{E}(Y) = \mathbb{E}(T^2) = \int_0^3 t^2 \frac{2}{9} t(3-t)\,dt = \frac{2}{9}\int_0^3 (3t^3 - t^4)\,dt = \frac{27}{10}.$$

(c) (7 points) Compute $\mathbb{V}\mathrm{ar}(Y)$.

Solution: We use

$$\mathbb{V}\mathrm{ar}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2.$$

First compute $\mathbb{E}(Y^2)$. Since $Y = T^2$, we have $Y^2 = T^4$, and the expectation formula gives

$$\mathbb{E}(Y^2) = \mathbb{E}(T^4) = \int_0^3 t^4 \frac{2}{9} t(3-t)\,dt = \frac{2}{9}\int_0^3 (3t^5 - t^6)\,dt = \frac{81}{7}.$$

Therefore

$$\mathbb{V}\mathrm{ar}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \frac{81}{7} - \left(\frac{27}{10}\right)^2 = \frac{81(100 - 3^2 \times 7)}{700} = \frac{81 \times 37}{700} = \frac{2997}{700}.$$

3. During a fixed time interval, a help desk receives a random number N of requests, where

$$N \sim \mathrm{Poi}(6).$$

Each request is independently classified as type A with probability $1/3$, and as type B with probability $2/3$. Let X be the number of type A requests, and let Y be the number of type B requests.

(a) (4 points) Find $\mathbb{P}(X = i, Y = j \mid N = i + j)$, for $i, j \in \{0, 1, 2, \ldots\}$.

Solution: Conditionally on $N = i + j$, the number of type A requests has binomial distribution with parameters $i + j$ and $1/3$. Hence

$$\mathbb{P}(X = i, Y = j \mid N = i + j) = \binom{i+j}{i} \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^j.$$

(b) (5 points) Find the joint probability mass function of X and Y.

Solution: The event $\{X = i, Y = j\}$ implies $N = i + j$. Therefore

$$\begin{aligned} \mathbb{P}(X = i, Y = j) &= \binom{i+j}{i} \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^j e^{-6} \frac{6^{i+j}}{(i+j)!} \\ &= e^{-6} \frac{2^i}{i!} \frac{4^j}{j!}, \qquad i, j \in \{0, 1, 2, \ldots\}. \end{aligned}$$

(c) (4 points) Find the marginal distributions of X and Y.

Solution: From the joint PMF,

$$\mathbb{P}(X = i) = \sum_{j=0}^{\infty} \mathbb{P}(X = i, Y = j) = e^{-2} \frac{2^i}{i!}, \qquad i = 0, 1, 2, \ldots,$$

so $X \sim \mathrm{Poi}(2)$. Similarly,

$$\mathbb{P}(Y = j) = \sum_{i=0}^{\infty} \mathbb{P}(X = i, Y = j) = e^{-4} \frac{4^j}{j!}, \qquad j = 0, 1, 2, \ldots,$$

so $Y \sim \mathrm{Poi}(4)$.

(d) (3 points) Are X and Y independent? Justify your answer.

Solution: Yes. For all $i, j \geq 0$,

$$\mathbb{P}(X = i, Y = j) = e^{-6} \frac{2^i}{i!} \frac{4^j}{j!} = \left(e^{-2} \frac{2^i}{i!}\right) \left(e^{-4} \frac{4^j}{j!}\right) = \mathbb{P}(X = i)\mathbb{P}(Y = j).$$

(e) (2 points) Compute $\mathbb{P}(N = 5 \mid X = 2)$.

Solution: Since $N = X + Y$, and X and Y are independent,

$$\mathbb{P}(N = 5 \mid X = 2) = \mathbb{P}(Y = 3 \mid X = 2) = \mathbb{P}(Y = 3) = e^{-4}\frac{4^3}{3!}.$$

4. A measuring device is used repeatedly to estimate an unknown quantity μ. The successive measurements $X_1, X_2, \ldots$ are independent and identically distributed. They satisfy

$$\mathbb{E}(X_i) = \mu, \qquad \mathbb{V}\text{ar}(X_i) \leq 4$$

for every i.

For $n \geq 1$, let

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

(a) (14 points) Using Chebyshev's inequality, prove that $\overline{X}_n \to \mu$ in probability and find an integer n large enough to guarantee

$$\mathbb{P}\left(\left|\overline{X}_n - \mu\right| < \frac{1}{2}\right) \geq 0.99.$$

Solution: By linearity of expectation,

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i) = \mu.$$

Since the measurements are independent, the variance of the sum is the sum of the variances:

$$\mathbb{V}\text{ar}(\overline{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{V}\text{ar}(X_i) \leq \frac{4}{n}.$$

Chebyshev's inequality says that, for any random variable Z with finite variance and any $\varepsilon > 0$,

$$\mathbb{P}(|Z - \mathbb{E}(Z)| \geq \varepsilon) \leq \frac{\mathbb{V}\text{ar}(Z)}{\varepsilon^2}.$$

Applying this to $Z = \overline{X}_n$, for every fixed $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\overline{X}_n - \mu\right| \geq \varepsilon\right) \leq \frac{\mathbb{V}\text{ar}(\overline{X}_n)}{\varepsilon^2} \leq \frac{4}{n\varepsilon^2} \longrightarrow 0.$$

By the definition of convergence in probability, this proves $\overline{X}_n \to \mu$ in probability.

With $\varepsilon = 1/2$, Chebyshev's inequality and the variance bound above give

$$\mathbb{P}\left(\left|\overline{X}_n - \mu\right| \geq \frac{1}{2}\right) \leq \frac{16}{n}.$$

To make this at most 0.01, it is enough to take $n \geq 1600$. Thus $n = 1600$ is large enough.

5. Three students arrive independently at random times during the same one-hour period. Let U_1, U_2, U_3 be their arrival times, measured in hours after the start of the period. Assume that U_1, U_2, U_3 are independent and uniformly distributed on $(0, 1)$.

Let

$$A = \min\{U_1, U_2, U_3\}, \qquad B = \max\{U_1, U_2, U_3\}, \qquad W = B - A.$$

Here A is the first arrival time, B is the last arrival time, and W is the time between the first and the last arrival.

(a) (8 points) Show that the joint density of (A, B) is

$$f_{A,B}(a, b) = \begin{cases} 6(b - a), & 0 < a < b < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Hint: You might want to first compute $F_{A,B}(a, b)$ for $0 < a < b < 1$.

Solution: By definition,

$$F_{A,B}(a, b) = \mathbb{P}(A \le a,\ B \le b).$$

For $0 < a < b < 1$, the event $\{B \le b\}$ means that all three arrival times are at most b. Inside this event, the complement of $\{A \le a\}$ is the event that all three arrival times are larger than a. Hence

$$\{A \le a,\ B \le b\} = \{U_1, U_2, U_3 \le b\} \setminus \{a < U_1, U_2, U_3 \le b\}.$$

Therefore, using independence and uniformity,

$$F_{A,B}(a, b) = b^3 - (b - a)^3, \qquad 0 < a < b < 1.$$

When a joint density exists, it is obtained from the joint CDF by

$$f_{A,B}(a, b) = \frac{\partial^2}{\partial a\, \partial b} F_{A,B}(a, b).$$

For $0 < a < b < 1$,

$$\frac{\partial}{\partial a} F_{A,B}(a, b) = 3(b - a)^2,$$

and hence

$$f_{A,B}(a, b) = \frac{\partial}{\partial b} 3(b - a)^2 = 6(b - a).$$

Thus

$$f_{A,B}(a, b) = \begin{cases} 6(b - a), & 0 < a < b < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(b) (8 points) Compute $\mathrm{Cov}(A, B)$.

Solution: We use the expectation formula for a joint density:

$$\mathbb{E}(g(A,B)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(a,b) f_{A,B}(a,b)\, db\, da.$$

Hence

$$\mathbb{E}(A) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a f_{A,B}(a,b)\, db\, da = \int_0^1 \int_a^1 6a(b-a)\, db\, da = \frac{1}{4},$$

and

$$\mathbb{E}(B) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b f_{A,B}(a,b)\, db\, da = \int_0^1 \int_a^1 6b(b-a)\, db\, da = \frac{3}{4}.$$

Also,

$$\mathbb{E}(AB) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ab f_{A,B}(a,b)\, db\, da = \int_0^1 \int_a^1 6ab(b-a)\, db\, da = \frac{1}{5}.$$

Therefore

$$\mathrm{Cov}(A,B) = \mathbb{E}(AB) - \mathbb{E}(A)\mathbb{E}(B) = \frac{1}{5} - \frac{1}{4}\cdot\frac{3}{4} = \frac{1}{80}.$$

(c) (2 points) Are A and B independent? Justify your answer.

Solution: If two random variables are independent, then their covariance is 0. Here $\mathrm{Cov}(A,B) = 1/80 \neq 0$, so A and B cannot be independent.

(d) (8 points) Find the cumulative distribution function and density of W.
Hint: You might want to first compute the joint density of (A,W).

Solution: Use the transformation formula for densities. With $w = b - a$, the inverse transformation is

$$(a,w) \mapsto (a,b) = (a, a+w),$$

and

$$f_{A,W}(a,w) = f_{A,B}(a,a+w)\left|\det\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}\right|.$$

The Jacobian determinant is 1, and the transformed support is $0 < a < 1-w,\ 0 < w < 1$. Hence

$$f_{A,W}(a,w) = f_{A,B}(a,a+w) = 6w, \qquad 0 < a < 1-w,\ 0 < w < 1,$$

and $f_{A,W}(a,w) = 0$ otherwise.
From a joint density, the marginal density is obtained by integrating out the other variable:

$$f_W(w) = \int_{-\infty}^{\infty} f_{A,W}(a,w)\, da = \int_0^{1-w} 6w\, da = 6w(1-w), \qquad 0 < w < 1.$$

The CDF is obtained from the density by

$$F_W(w) = \int_{-\infty}^{w} f_W(u)\,du.$$

Thus

$$F_W(w) = \begin{cases} 0, & w < 0, \\ 3w^2 - 2w^3, & 0 \le w \le 1, \\ 1, & w > 1. \end{cases}$$

The density is $f_W(w) = 6w(1-w)$ for $0 < w < 1$, and 0 otherwise.