American University of Beirut

MATH/STAT 233: Probability

2022–2023 Spring

Final exam (Thursday, May 4, 17:30 – Nicely 214)

How it works: You receive the questions in advance. At the beginning of the exam, you declare which of the questions you have managed to solve. For each question, one student (from those who have solved it) will be called to show their solution on the board.

What is permitted and what is not permitted: Before the exam, you can (and are encouraged to) discuss the problems with the other students in class but not with external people. Getting help from internet forums is also not permitted. Feel free to use your class notes, books, computers, or Wikipedia.

Problem 1 (Two events)**.** Prove that if A and B are two events such that $A \cup B = \Omega$, then

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B) - \mathbb{P}(A^c)\,\mathbb{P}(B^c)\ .$$

Problem 2 (Independence)**.** Let X, Y and Z be three random variables. Prove or disprove: If X is independent of Z, and Y is independent of Z, then $X + Y$ is also independent of Z.

Problem 3 (A curious inequality)**.**

(a) Argue that if X and Y are independent and identically distributed random variables, then $\mathbb{P}(X \leq Y) \geq 1/2$.

(bonus) (b) Would the latter inequality still hold if we removed the independence assumption?

(c) Prove that if X_1, X_2 and Y are i.i.d. random variables, then

$$\mathbb{P}\left(\left|\frac{X_1 + X_2}{2}\right| \leq |Y|\right) \geq 1/4\ .$$

[*Hints*: (i) Triangle inequality, (ii) In order to have $a + b \leq 2c$, it is sufficient that $a \leq c$ and $b \leq c$, (iii) Conditioning on Y, (iv) The inequality $\mathbb{E}[Z^2] \geq \mathbb{E}[Z]^2$.]

(d) Generalize the latter argument to obtain a similar inequality for i.i.d. random variables $X_1, X_2, \ldots, X_n, Y$. [*Hint*: The inequality $\mathbb{E}[Z^n] \geq \mathbb{E}[Z]^n$ (for $n \geq 1$) is an instance of Jensen's inequality.]

Problem 4 (Generating functions)**.**

(a) Use generating functions to show that the sum of two independent Poisson random variables is again a Poisson random variable.

(b) Let X and Y be independent random variables. Prove that if X and $X + Y$ are Poisson, then so is Y.

Problem 5 (Rounding error)**.** We would like to find the sum of 100 numbers. Student #1 adds all the numbers and then rounds the result to the nearest integer. Student #2 rounds each number separately and then adds them up. We assume that the decimal parts of the 100 numbers are independent, each distributed uniformly over $[0, 1)$.

(a) Which of the two computation methods provides a better estimate?

(b) Use the central limit theorem to estimate the probability that the value obtained by the second student is off by more than 3. [*Hint*: Let X_k denote the k'th number. Observe that the rounding error $E_k := \mathrm{round}(X_k) - X_k$ in the k'th number is uniformly distributed over $[-1/2, 1/2]$.]

(c) Use the central limit theorem to estimate the probability that the values obtained by the two students coincide. [*Hint*: Observe that $S_2 = S_1$ if and only if the total error of the second student $E := E_1 + E_2 + \cdots + E_{100}$ is in the interval $[-1/2, 1/2]$.]

Feel free to use a computer/calculator for computations concerning the normal distribution.

Problem 6 (Expected minimum of Pareto RVs). Let $\alpha > 0$ be a constant. Let X and Y be independent random variables with the same probability density function given by

$$f(x) = \begin{cases} \dfrac{\alpha}{x^{\alpha+1}} & \text{if } x > 1, \\ 0 & \text{if } x < 1. \end{cases}$$

Find $\mathbb{E}[\min(X, Y)]$. [*Hint*: Use $\mathbb{E}[Z] = \int_0^\infty \mathbb{P}(Z > z)\, \mathrm{d}z$ if applicable.]

Problem 7 (Joint distribution). Two random variables X and Y have joint density

$$f(x, y) = \begin{cases} \mathrm{e}^{-y} & \text{if } 0 < x < y < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal distributions of X and Y.

(b) Find the density function of $Y - X$. [*Hint*: Start by finding its cdf.]

Problem 8 (Random sum). Let $X_1, X_2, \ldots$ be independent random variables with the same distribution, and define $S_0 \coloneqq 0$ and $S_n \coloneqq X_1 + X_2 + \cdots X_n$ for $n = 1, 2, \ldots$. Let N be a non-negative integer-valued random variable that is independent of the X_i's. Consider the random variable S_N.

(a) Prove that $\mathbb{E}[S_N] = \mathbb{E}[X_1]\, \mathbb{E}[N]$, provided that the expected values on the right-hand side exist. [*Hint*: Condition on N as we did in class.]

(b) Prove that $\mathbb{V}\mathrm{ar}[S_N] = (\mathbb{E}[X_1])^2\, \mathbb{V}\mathrm{ar}[N] + \mathbb{E}[N]\, \mathbb{V}\mathrm{ar}[X_1]$ provided the expected values and the variances on the right-hand side exist. [*Hint*: Write $\mathbb{V}\mathrm{ar}[S_N] = \mathbb{E}[S_N^2] - \mathbb{E}[S_N]^2$ and proceed as in the previous part.]

(c) The number of road accidents in Beirut during one month can be modelled as a Poisson random variable with parameter μ. Each accident has a probability p of being fatal, and the fatality of different accidents can be assumed to be independent. Show that the number of fatal accidents during one month is also Poisson-distributed and identify its parameter.

(d) Let X_1 be the arrival time of the first customer at a store, and for $n \geq 2$, let X_n be the inter-arrival time between the $(n-1)$'st and the n'th customers. We assume that $X_1, X_2, \ldots$ are independent, exponentially-distributed random variables with the same rate λ. Each customer has a probability p of making a purchase, independently of the other customers. Suppose that the first purchase is made by the N'th customer. Prove that S_N, the arrival time of the first purchasing customer, is exponential with parameter $p\lambda$. [*Hint*: Start by writing the moment generating function of S_N in terms of the the moment generating function of X_1 and the probability generating function of N. Use the fact that the moment generating function of a RV uniquely determines its distribution.]