Семинар - Теория вероятностей - 2021/22

1 - комбинаторика и вероятность: Задачи

З.1. Из колоды (52 карты) вынимают 4 карты.

(1a) Какова вероятность, что все 4 карты — пики?

(1b) Какова вероятность, что 3 карты — пики, а одна — черви ?

З.2. Колоду из 52 карт раздают на 4 игроков. Один из игроков объявляет, что у него есть туз.

(2a) Какова вероятность, что у него есть еще хотя бы один туз?

(2b) Какова вероятность, что у него есть еще хотя бы один туз, если у него уже есть туз пик?

З.3. Человек несколько раз подбрасывает монету. Он останавливается, как только он получает последовательность из n орлов или последовательность, из 1 решки и $(n-1)$ орлов, идущих подряд.

(3a) Какова вероятность того, что он остановится?

(3b) Какова вероятность, что он остановится на последовательности из n орлов?

З.4. На плоскости отметили точки с координатами (x, y), где $x, y \in \{1, 2, \ldots, N\}$ (с равной вероятностью). Наугад выбирают точку (x, y) из отмеченных. Пусть p_N вероятность того, что она лежит в круге радиуса N. Вычислите $\lim_{N\to\infty} p_N$.

З.5. Так работает (одномерная) линия передачи сотовой сети. n антенн выстроены в линию на равном расстоянии, каждая повторяет сигнал, полученный от соседней антенны. Сигнал может преодолевать расстояние двух антенн, поэтому передача работает, если нет двух последовательных неисправных антенн. Предположим, что $m < n$ антенн неисправны, но мы не знаем их положения. Какова вероятность того, что трансмиссия сработает?
I I X I I I X I = работает
I I X X I I I I = не работает

З.6. 200 студентов, изучающих теорию вероятности, разделены на три группы по 50, 50 и 100 студентов в каждой, чтобы посещать семинары в классах с соответствующей вместимостью.

(6a) Сколькими способами можно сформировать эти группы студентов?

Александра и Светлана - хорошие друзья, и во время семинара хотели бы находиться в одной комнате. Но Алексея терпеть не могут, и они надеются, что его переведут в другую комнату.

(6b) Какова вероятность того, что их пожелания будут выполнены?

З.7. Человек одновременно купил две коробки спичек и положил их в карман. После этого каждый раз, когда ему нужно было зажечь спичку, он доставал наудачу ту или иную коробку. Через некоторое время, вытащив одну из коробок, человек обнаружил, что она пуста. Какова вероятность, что в другой коробке в этот момент находилось k спичек, если число спичек в новой коробке равно n?[1]

З.8. Электричка состоит из n вагонов. Каждый из k пассажиров выбирает вагон наудачу. Какова вероятность, что в каждом вагоне будет хотя бы один пассажир? Какова вероятность, что будут заняты ровно r вагонов?

З.9. Урна пуста. Каждую секунду человек кладет в урну 100 шариков и случайным образом извлекает один шарик. (например, после 10 секунд будут 990 шариков).

(9a) Какова вероятность, что шарик, который был положен первую секунду, останется в урне не менее n секунд?

(9b) Какова вероятность, что шарик из пункта (a) останется в урне навсегда?

(9c) Какова вероятность, что хотя бы один шарик останется в урне навсегда?

(9d) Это парадокс?

З.10. E и F — два конечных множества. Строим функцию $f\colon E \to F$ следующим образом. Для каждого элемента $x \in E$ мы выбираем $f(x) = y$ с вероятностью p_y. Конечно $\sum_{y\in F} p_y = 1$. Какова вероятность того, что изображение функции f содержит заданное подмножество $B \subset F$?

[1] Классический анекдот к этому упражнению состоит в том, что этот человек - Банах, у него в каждой руке по бутылке водки, и он пьет из них беспорядочно.

2 - Условная вероятность: Задачи

З.11. Из множества $\{1, 2, ..., n\}$ без возвращения по очереди выбирают три различных числа. Найдите условную вероятность того, что третье число лежит между первым и вторым при условии, что первое число меньше второго.

З.12. Пусть $n \geq 2$. Из множества $\{1, 2, \ldots, n\}$ случайным образом выбирается одно число. Событие A состоит в том, что выбранное число делится на 2, а событие B — в том, что выбранное число делится на 7. Найдите все такие n, что события A и B независимы.

З.13. В урне находятся u_1 шаров цвета 1, u_2 шаров цвета 2, $\ldots$, u_R шаров цвета R. Выполняем n случайных извлечений, и сразу после каждого извлечения шар возвращается в урну вместе с m других шаров того же цвета ($m \geq -1$ и $n \leq u_1 + \ldots u_R$, если $m = -1$).

(13a) Какова вероятность того, что на i-м извлечении будет вынут шар цвета r?

(13b) Какова вероятность того, что среди n выбранных шариков цвет 1 встречается j_1 раз, цвет 2 встречется j_2 раз, ..., цвет R встречается j_R раз?

(13c) Пусть $u_1 = u_2 = \ldots = u_R = m$. Вычислите предыдущую вероятность в следующих случаях: $m = -1$ (извлечение шаров без возвращения); $m = 0$ (извлечение шаров с возвращением).

З.14. События A, B, C *попарно* независимы и равновероятны, $A \cap B \cap C = \emptyset$. Найти максимально возможное значение $P(A)$.

З.15. Пусть p_n обозначает вероятность того, что за n подбрасываний симметричной монеты не случится выпадения трех орлов подряд. Найдите рекуррентное соотношение для p_n.

З.16. По кругу стоят 15000 солдат из 1000 групп по 15 человек. Докажите, что в каждой группе можно выбрать старосту так, что рядом не окажется двух старост.

З.17. Агент Д. следит за передвижениями президента некоторой компании. Известно, что президент бывает в офисе с вероятностью 60%, а на даче с вероятностью 40%. У агента Д. есть два осведомителя, причем первый ошибается с вероятностью 20%, а второй – с вероятностью 10%. Первый осведомитель утверждает, что президент компании в офисе, а второй осведомитель утверждает, что он на даче. Где находится президент?

З.18. Принцесса выбирает жениха. Свататься приехали N женихов. Про любых двух из них принцесса может сказать, кто лучше, а кто хуже. Все женихи образуют упорядоченное множество, т.е. если A лучше B и B лучше C, то A лучше C. Женихи заходят к принцессе по очереди в случайном порядке. Если невеста отказывает жениху, то он сразу уезжает. Если принцесса выбирает жениха, то на этом просмотр женихов заканчивается. Как действовать принцессе, чтобы с наибольшей вероятностью выбрать лучшего жениха?

З.19. Монета с вероятностью выпадения орла p подбрасывается N раз. Найдите вероятность того, что

(19a) выпало ровно три орла;

(19b) число выпавших орлов четно;

(19c) выпало орлов меньше половины.

З.20. Алиса и Боб играют в следующую игру. Правильная монета подбрасывается до тех пор, пока не встретится комбинация 110 или 100. Алиса выигрывает, если первой появилась комбинация 110, а Боб в случае, когда первой появилась комбинация 100. Кто будет выигрывать чаще?

З.21. X и Y договорились встретиться в промежуток времени с 12:00 до 13:00, причем каждый из них готов ждать ровно 30 минут. Какова вероятность встречи? Какова вероятность того, что они встретились и X не ждал Y? Какова вероятность того, что они пришли одновременно?

З.22. Трое загадывают по числу из отрезка $[0, 1]$. Какова вероятность того, что существует треугольник с такими сторонами?

3 - Случайные величины: Задачи

3.23. В коробке 7 красных и 5 белых шаров. Случайным образом из коробки вынимают два шара. Найдите математическое ожидание и дисперсию числа вынутых красных шаров. Изменится ли ответ, если вынимать шары следующим образом: вытащили первый шар и положили обратно, а затем вытащили второй шар?

3.24. Случайная величина X имеет пуассоновское распределение с параметром λ_1, случайная величина Y распределена экспоненциально с параметром λ_2, причем X и Y независимы. Найдите математическое ожидание и дисперсию случайных величин $X+Y$, XY, $\max(X,Y)$.

3.25. На матфаке два преподавателя читают ТерВер. Первый уже знает, что в его классе n студентов. Другой, однако, пропустил первое занятие, поэтому он считает, что количество N его студентов – случайная величина с мат. ожиданием равным n.

Планируется провести экзамен, вероятность успешной сдачи которого равна $p \in (0,1)$ для каждого студента (независимо от других студентов и от N). В каком классе наибольшее ожидаемое число тех студентов, которые успешно сдадут экзамен? В каком классе набольшая дисперсия числа студентов, которые сдадут экзамен?

3.26. $n = 100$ писем разложили по n конвертам, на которых уже были написаны адреса, случайным образом. Найдите математическое ожидание и дисперсю количества писем, лежащих в конвертах с правильными адресами.

3.27. Монету, для которой вероятность выпадения орла равна p, бросают до первого выпадения орла. Найдите математическое ожидание и дисперсию числа бросаний в случае

(27a) число бросаний не более N. Мы полагаем, что на N-ом броске выпадает орел вне зависимости от настоящего исхода.

(27b) число бросаний неограничено.

3.28. Пусть $X \sim \mathcal{N}(0,1)$ и $Y \sim \text{Poisson}(1)$ независимые случайные величины. Нарисуйте функцию распределения с.в. $X+Y$.

3.29. $S^1 = \mathbb{R}/\mathbb{Z}$ - окружность длины 1. Пусть $I_0 = [0, 1/n] \subset S^1$ и $I_k = \frac{k}{n} + I_0 = [k/n, (k+1)/n]$, $k = 1, \dots, n-1$. тогда $|I_0| = 1/n$ и $|\cup_{i=0}^{n-1} I_k| = 1$. Таким образом, мы построили n сдвигов множества I_0 таких, что их объединение имеет полную меру.

Однако, в общем случае мы имеем иную картину. Для $E \subset S^1$ и $x \in S^1$ положим $x + E := \{x + y, y \in E\}$. Пусть множество E измеримо и $n \geq 1$. Докажите, что существуют $x_1, \dots, x_n$ такие, что $|\cup_{i=1}^{n} (x_i + E)| \geq 1 - (1 - |E|)^n$.

3.30. Пусть X - случайная величина. Вещественное число $m \in \mathbb{R}$ называется *медианой* случайной величины X, если $\mathbb{P}(X \leq m) \geq 1/2$ и $\mathbb{P}(X \geq m) \geq 1/2$. Докажите следующие утверждения.

(30a) Если $X \in L^1$, то медиана m минимизирует функцию $\phi_1(r) = \mathbb{M}[|X - r|]$.

(30b) Если $X \in L^2$, то математическое ожидание $\mathbb{M}[X]$ минимизирует функцию

$$\phi_2(r) = \mathbb{M}[|X - r|^2]$$

(30c) Используйте (a) и неравенство Йенсена, чтобы доказать, что $|m - \mathbb{M}[X]|^2 \leq \mathbb{D}(X)$.

3.31. Мы видели, что если $X \sim \mathcal{N}(0,1)$, то $\mathbb{M}[f'(X)] = \mathbb{M}[Xf(X)]$. Можете ли вы обобщить эту формулу на случай, когда X — $\mathbb{R}^d$-значная случайная величина с $X \sim \mathcal{N}(0, S)$?

3.32. Пусть $X \sim \mathcal{N}(0, \sigma^2)$. Вычислите $\mathbb{M}[X^n]$ для $n \in \mathbb{N}$.

3.33. Пусть E - упорядоченное измеримое пространство. Пусть $X \colon \Omega \to E$ - случайная величина. Пусть $f, g \colon E \to \mathbb{R}$ - измеримые монотонные функции. Докажите, что

$$\mathbb{M}[f(X)\, g(X)] \geq \mathbb{M}[f(X)]\mathbb{M}[g(X)]$$

3.34. Напомним неравенство Чебышева:

$$\mathbb{P}(|X - \mathbb{M}[X]| \geq c) \leq D(X)/c^2 \tag{1}$$

Докажите, что для каждого $a, c \geq 0$ т.ч. $a \leq c^2$, существует такая случайная величина X, т.ч. $D(X) = a$ и в (1) будет равенство $\mathbb{P}(|X - \mathbb{M}[X]| \geq c) = D(X)/c^2$.

3.35. Пусть на полосу бесконечной длины и единичной ширины на плоскости случайным образом бросается игла единичной длины. Какова вероятность того, что игла пересечет хотя бы одну из линий, образущих полосу?

3.36. Точка (x, y) выбиратся из квадрата $[0, 1] \times [0, 1]$ согласно равномерному распределению. Найдите распределение случайной валичины $x^2; x/(x + y); x^2 + y^2$, $\min(x, y)$, $\max(x, y)$.

3.37. Пусть вектор (ξ, η) равномерно распределен на $[0, 1] \times [0, 1]$. Найдите распределение вектора (X, Y) где

$$X = \sqrt{-2\log(\xi)}\cos(2\pi\eta)$$
$$Y = \sqrt{-2\log(\xi)}\sin(2\pi\eta)$$

3.38. Пусть (X, Y) имеет равномерное распределение в $(x, y) : x^2 + y^2 \leq 1$. Являются ли величины X и Y независимыми?

3.39. $X_1, X_2, \ldots$ — независимые случайные величины, с одинаковым распределением $\exp(\lambda)$. Пусть $Y_n := \sum_{i=1}^{n} X_i$ и $N_t := \inf\{n \geq 0 \,:\, Y_{n+1} > t\}$, $t > 0$.

(39a) Докажите что распределение Y имеет плотность $\varrho_n(y) := e^{-\lambda y}\frac{\lambda^n y^{n-1}}{(n-1)!}\mathbf{1}_{y\geq 0}$.

(39b) Докажите что $\mathbb{P}(N_t = k) = e^{-\lambda t}(\lambda t)^k/k!$. (это означает, что $N_t \sim \mathrm{Poisson}(\lambda t)$).

3.40. Пусть X - (вещественная) случайная величина. Предположим, что существует $r > 0$ т.ч. $\mathbb{M}[e^{rX}] < \infty$. Определим функцию

$$\psi(t) := \log \mathbb{M}[e^{tX}], \qquad t \in [0, r) \tag{2}$$

(40a) Докажите, что ψ — выпуклая функция, т.е.

$$\psi((1 - \alpha)t + \alpha s) \leq (1 - \alpha)\psi(t) + \alpha\psi(s), \qquad \alpha \in [0, 1]$$

(40b) Пусть ψ^* — выпуклое сопряжение ψ (преобразование Лежандра):

$$\psi^*(x) = \sup_{t\in[0,r)} (t\,x - \psi(t)), \qquad x \in \mathbb{R} \tag{3}$$

Докажите, что $\psi^* \geq 0$ и, если $\mathbb{M}[X] > -\infty$ и $x \leq \mathbb{M}[X]$, то $\psi^*(x) = 0$.

(40c) Докажите что

$$\mathbb{P}(X \geq x) \leq e^{-\psi^*(x)}, \qquad x \in \mathbb{R}$$

Выведите отсюда, что если $X_1, \ldots, X_N$ — независимые, одинаково распределенные случайные величины, и для каждой из них ψ определяется как в (2), то

$$\mathbb{P}\left(\tfrac{1}{N}\sum_{j=1}^{N} X_j \geq x\right) \leq e^{-N\psi^*(x)}$$

(40d) Пусть $\bar{r} := \sup\left\{r \geq 0 \,:\, \mathbb{M}[e^{r|X|}] < +\infty\right\}$. Докажите, что ψ продолжается до аналитической функции в комплексном круге $\{|z| < \bar{r}\}$.

(40e) Приведите пример случайной величины т.ч. $\bar{r} > 0$ и $\lim_{t\uparrow\bar{r}} \psi(t) = +\infty$. Приведите пример случайной величины т.ч. $\bar{r} > 0$ и $\lim_{t\uparrow\bar{r}} \psi(t) < +\infty$.

3.41. Существуют ли независимые непостоянные случайные величины X и Y такие, что $X^2 + Y^2 = 1$?

3.42. Опишите все случайные величины X такие, что X и $\sin(X)$ независимы.

3.43. Лампочка работает случайное время T, которое экспоненциально распределено с параметром $\lambda = 10^{-6}$ секунд $^{-1}$. Мощность w_t, которую она излучает в момент t, случайна, независима от T, и равномерно распределенна на интервале $[a, b] = [39, 41]$ ватт. Рассчитайте мат. ожидание энергии, потребляемой лампочкой в течение срока ее службы. Здесь мы предполагаем, что функция $t \mapsto w_t(\omega)$ непрерывна для каждого ω.

4 - Характеристические функции: Задачи

З.44. Пусть случайные величины ξ и η имеют стандартное нормальное распределение, а случайная величина α – равномерное распределение на отрезке $[0, 2\pi]$. Найдите распределение случайной величины $\xi\cos(\alpha) + \eta\sin(\alpha)$, если известно, что ξ, η и α независимы.

З.45. Пусть X, Y, Z – независимые случайные величины со стандартным нормальным распределением. Найдите распределение случайной величины $\frac{X+ZY}{\sqrt{1+Z^2}}$.

З.46. Докажите, что сумма независимых пуассоновских случайных величин является пуассоновской. Приведите контрпример для случая зависимых случайных величин величин.

З.47. Объясните, почему следующие функции не являются характеристическими:

$$\sin(t), \qquad \cos(t^2), \qquad 1 + \sin(t), \qquad e^{-t^4}.$$

З.48. Приведите пример некоррелированных гауссовских случайных величин X и Y (то есть $\mathbb{M}[XY] = \mathbb{M}[X]\mathbb{M}[Y]$) таких, что вектор (X, Y) не является гауссовским.

З.49. Пусть случайная величина X имеет распределение Коши (напомним, что распределение Коши имеет плотность $\varrho(x) = \frac{1}{\pi(1+x^2)}$). Найдите характеристическую функцию X. Почему она не дифференцируема в нуле?

З.50. Пушка расположена на расстоянии 1 от траншеи противника, представляющей из себя бесконечную прямую линию. Артиллерист производит выстрел в случайном направлении под углом $\theta \in [-\pi/2, \pi/2]$. Зафиксируем систему координат таким образом, что $\theta = 0$ соответствует направлению, перпендикулярному к траншее, а точка $x = 0$ – ближайшей к пушке точке траншеи. Если ядро попадает в точку с координатой x, она наносит случайный ущерб V_x, который распределен согласно экспоненциальному закону $\exp(\lambda_x)$, где $\lambda_x = e^{|x|}/(1+x^2)$. Более того, $\mathbb{M}[V_xV_y] - \mathbb{M}[V_x]\mathbb{M}[V_y] = \frac{e^{-|x-y|}}{\lambda_x\lambda_y}$, где $c \geq 0$ — некоторая константа (тут опечатка, c отсутствует в формуле) (то есть корреляция ущерба, нанесенного точкам x и y, падает с увеличением расстояния между этими точками).

(50a) Вычислите математическое ожидание и дисперсию ущерба от одного выстрела.

Пусть было произведено два выстрела, направления которых были выбраны независимо.

(50b) Вычислите математическое ожидание и дисперсию суммарного ущерба от двух выстрелов.

5 - Условное мат.ожидание: Задачи

З.51. Пусть μ и ν - положительные σ-конечные меры на измеримых пространствах E и F. Пусть X и Y - случайные величины, принимающие значения в E и F соответственно, и такие, что закон $\theta \in \mathcal{P}(E \times F)$ пары (X, Y) имеет плотность

$$d\theta(x, y) = \varrho(x, y) d\mu(x) d\nu(y)$$

относительно $\mu \otimes \nu$. Для ограниченной измеримой функции $f \colon E \times E' \to \mathbb{R}$, запишете $\mathbb{M}[f(X, Y)|Y]$ как функцию, зависящую от ϱ, μ, ν и f.

(здесь под законом понимается образ меры $\gamma_1 \otimes \gamma_2$ на вероятностном пространстве $(A \times B, \mathcal{F}_A \times \mathcal{F}_B, \gamma_1 \otimes \gamma_2)$ под действием отображения $(\omega_1, \omega_2) \ni (A, B) \mapsto (X(\omega_1), Y(\omega_2)))$

З.52. Пусть распределение $p(X, Y)$ пары случайных величин $(X, Y) \in \mathbb{R}^2$ совместное гауссовское. Является ли распредение $p(Y \mid X)$ гауссовским?

З.53. Модель роста человека в семье устроена следующим образом: рост H_i члена i семейства - это $H_i = X + Y_i$, где X и Y_i - независимые гауссовские с средним значением $m = 85\, cm$ и дисперсией $\sigma^2 = 25\, cm^2$. X одинаков для всех членов семьи и учитывает генетику. (Y_i) независимы для разных i.

Рост Арнольда 180 см. Найдите математическое ожидание роста его брата Бориса.

З.54. $(\Omega, \mathfrak{F}, P)$ - вероятностное пространство, $X \in L_2(\Omega, \mathfrak{F}, \mathbb{P})$ - величина, $\mathfrak{G} \subset \mathfrak{F}$ - σ-алгебра. Найдите случайную величину $\bar{Y}$, минимизирующую $\mathbb{M}[(X - Y)^2]$ в классе $Y \in L_2(\Omega, \mathfrak{G}, \mathbb{P})$.

6 - Сходимость: Задачи

З.55. Пусть $X_n \sim \mathcal{N}(0, 1/n)$ и $\mathbb{P}(X = 0) = 1$. Для каких $-\infty < a \le b < \infty$ выполнено соотношение $\lim_n \mathbb{P}(X_n \in [a, b]) = \mathbb{P}(X \in [a, b])$?

З.56. Пусть $X_n \sim \text{Bernoulli}(1/n)$. Постройте вероятностное пространство, в котором определены все $X_n, n \ge 1$ и $\mathbb{P}(\lim_n X_n = 0) = 1$. Постройте вероятностное пространство, в котором последнее равенство не выполняется. Верен ли второй пример, если $X_n \sim \text{Bernoulli}(2^{-n})$?

З.57. Пусть $\Omega = [0, 1]$, $\mathbb{P}$ — мера Лебега. Для любого $n \in \mathbb{N}$ определим $h_n, k_n \in \mathbb{N}$ такие, что $n = 2^{h_n} + k_n$ и $k_n < 2^{h_n}$ (h_n и k_n определены однозначно). Пусть $X_n \colon \Omega \to [0, 1]$

$$X_n(\omega) = \begin{cases} 1 & \text{если } \omega \in [2^{-h_n} k_n, 2^{-h_n}(k_n + 1)) \\ 0 & \text{в противном случае} \end{cases}$$

Верно ли, что 1) $\mathbb{P}(\lim_n X_n = 0)$; 2) $\lim_n \mathbb{P}(X_n > \varepsilon) = 0$ для всех $\varepsilon \in (0, 1)$; 3) $\sum_n \mathbb{P}(X_n > \varepsilon) < \infty$?

З.58. Пусть $p \ge 1$ и $c > 0$. Найдите такую последовательность (X_n) случайных величин, что выполнено

(58a) $\lim_n \mathbb{M}[|X_n|^q] = 0$, если $q < p$, $\lim_n \mathbb{M}[|X_n|^p] = c$ и $\lim_n \mathbb{M}[|X_n|^q] = +\infty$, если $q > p$.
(58b) $\lim_n \mathbb{M}[|X_n|^q] = 0$, если $q \le p$ и $\lim_n \mathbb{M}[|X_n|^q] = +\infty$, если $q > p$.
(58c) $\lim_n \mathbb{M}[|X_n|^q] = 0$, если $q < p$ и $\lim_n \mathbb{M}[|X_n|^q] = +\infty$, если $q \ge p$.

З.59. Пусть $V(x) = (x^2 - 1)^2$, и X_n – абсолютно непрерывная случайная величина с плотностью $\varrho_n(x) = \exp(-nV(x))/Z_n$. Найдите $\lim_n \mathbb{D}[3X_n^2 + \cos(2\pi X_n)]$. Указание: найдите $\lim \mathbb{M}[f(X_n)]$, где f — непрерывная ограниченная функция.

З.60. $X_n \sim \text{Binomial}(n, p_n)$. Принимаите $np_n \to \lambda > 0$. найдите $X = \lim_{n\to\infty} X_n$ по распределению.

З.61. Пусть (X_j) - последовательность независимых случайных величин, $X_j \sim$Bernoulli(p_j); пусть (a_j) – последовательность действительных чисел такая, что $\sup_j |p_j a_j| < \infty$. Найдите, если существует,

$$\lim_n \tfrac{1}{n} \sum_{j=1}^{n} a_j X_j,$$

по распределению.

З.62. Дана последовательность независимых гауссовских случайных величин $X_j \sim \mathcal{N}(0, 1)$. Верно ли, что последовательность $Z_n := (X_1 + \ldots + X_n)/\sqrt{n}$ сходится по вероятности?

З.63. Пусть X – абсолютно непрерывная случайная величина и $X_n \to X$ по распределению. Докажите, что

$$\lim_n \sup_I |\mathbb{P}(X_n \in I) - \mathbb{P}(X \in I)| = 0$$

где I – интервалы.

З.64. Игральную кость бросают $n = 42{*}10^7$ раз. Оцените вероятность того, что сумма выпавших очков превышает $l = 1470070000$.

З.65. Аренда суперкомпьютера Фугаку стоит очень (очень) дорого, ее оплата почасовая и вносится заранее. Вы хотите провести симуляцию сворачивания белка. Каждую секунду суперкомпьютер может генерировать (виртуальные) 1000 белков и проверять их стабильность.

Каждый белок стабилен с вероятностью $p = 10^{-5}$. Вам нужно найти как минимум 100 стабильных белков с вероятностью 95%. На сколько часов вы будете арендовать Фугаку?

З.66. Пусть X_n – случайная величина, $\mathbb{P}(X_n \in \mathbb{N}) = 1$. Пусть $p_{n,k} = \mathbb{P}(X_n = k)$, $k \in \mathbb{N}$. Предположим, что для каждого k существует предел $\lim_n p_{n,k}$. Сходится ли X_n по распределению?

1 - комбинаторика и вероятность: Решения

P.1. Recall that a French deck (poker deck) of 52 cards is composed by 4 *colors* (or *suits*), each one counting 13 *kinds* (or *values*).

(1a) There are $\binom{52}{4}$ ways to sample 4 cards out of 52. There are $\binom{13}{4}$ to sample black cards. So the probability is $\binom{13}{4}/\binom{52}{4} = \frac{13!48!}{52!9!} = \frac{13*12*11*10}{52*51*50*49} \simeq 0.0026$.

(1b) There are $\binom{13}{3}\binom{13}{1}$ ways to choose 3 black and one red card. So the answer is $\binom{13}{3}\binom{13}{1}/\binom{52}{4} \simeq 0.034$.

P.2. This can be a bit surprising, because why should the suit of the ace change the probability? Of course it does, having the ace of spades is not as easy as having *any* ace, and has an higher correlation with having any other ace. Let us calculate it quantitatively.

(2a) There are $\binom{52}{13}$ possible 13-cards hands to give to a player. Out of those many, $\binom{48}{13}$ contain no aces (since there are 4 aces in the deck). Thus $\binom{52}{13} - \binom{48}{13}$ are the hands containing at least one ace. This could be also calculated as

$$\underbrace{\binom{4}{1} * \binom{48}{12}}_{\text{exactly one ace}} + \underbrace{\binom{4}{2} * \binom{48}{11}}_{\text{exactly two aces}} + \underbrace{\binom{4}{3} * \binom{48}{10}}_{\text{exactly three}} + \underbrace{\binom{4}{4} * \binom{48}{9}}_{\text{exactly four}}$$

Similarly the number of hands with at least two aces is $\binom{52}{13} - \binom{48}{13} - \binom{4}{1} * \binom{48}{12}$. Thus the probability of having at least 2 aces, given that one has at least 1 ace is

$$\frac{\binom{52}{13} - \binom{48}{13} - \binom{4}{1}\binom{48}{12}}{\binom{52}{13} - \binom{48}{13}} \simeq 0.37 \tag{4}$$

(2b) Consider now the case where the player states that he has the ace of spades. Then there are $\binom{51}{12}$ (which also equals $\binom{52}{13} - \binom{51}{13}$) hands containing such an ace, and $\binom{48}{12}$ hands containing no other ace. Thus the probability of having at least another ace is

$$\frac{\binom{51}{12} - \binom{48}{12}}{\binom{51}{12}} \simeq 0.56$$

Notice that this value is larger than the one in formula (4).

One could have also reasoned as follows (this second intuitive solution is however harder to generalize in other cases). The probability of not having any other ace but the ace of spades, equals the probability of turning 12 cards (out of the 51 that are not the ace of spades) in a row, without getting any other ace. Since this probability is then intuitively calculated as $\frac{48}{51} \cdots \frac{37}{40}$, the required probability is

$$1 - \frac{48!39!}{36!51!} = 1 - \frac{\binom{48}{12}}{\binom{51}{12}}$$

giving the same result as above.

P.3. There is only one way he can see a sequence of n consecutive heads, before 1 tail and $n-1$ heads: this sequence appears right at the beginning. So the required probability is 2^{-n}. In general, given two sequences of heads and tails (or 0 and 1), it is hard to calculate exactly the probability that one appears before the other. Certainly, it does not depend only on the length of the sequence, as this exercise readily shows.

P.4. The required probability coincides with the following one (just looking at the same picture from distance N). Sample x and y uniformly and independently in $\{1/N, 2/N, \ldots, 1\}$, and let P_N be the probability that (x, y) falls with the (quadrant) circle of radius 1. As $N \to \infty$, we can simply use the fact that the circle is Riemann-measurable to argue that $p_N \to \frac{\text{measure of the circle quadrant}}{\text{measure of } [0,1]^2} = \pi/4$, as indeed squares with sides $1/N$ on $[0,1]^2$ provide a Riemannian mesh. The same approach would work for any Riemman-measurable subset of $[0,1]^2$.

On the other hand, the sharp asymptotic behavior of the quantity $p_N - \pi/4$ is quite a hard problem. Several celebrated mathematicians worked on the problem, including Gauss, Hardy and E.Landau.

P.5. Mark a I for functional antenna, and a X for a non-functional one. To identify a configuration of working-nonworking antennas, it is enough to give the position of the m X, out of the n possible positions. So we have $\binom{n}{m}$ equally likely possibilities.

How many of them do not have two consecutive X? To calculate, draw the following picture:
_ I _ I _ I ... _ I _

where there are $(n-m)$ I's. To produce a functional configuration, we can put the m non-functional antennas in any of the $n-m+1$ free spots marked with _. So the probability of having a working network is

$$\frac{\binom{n-m+1}{m}}{\binom{n}{m}}$$

which of course means 0 for $2m > n+1$.

P.6. This simple exercise is just to recall you that sometimes there are several different *natural* probability spaces where one can calculate.

(6a) There are $\binom{200}{50,50,100}$ ways to arrange the students in the three classrooms. However, if we consider the ways to divide the students in two groups of 50 and one group of 100, then we have $\binom{200}{50,50,100}/2$ arrangements.

(6b) For instance there are $\binom{197}{48,50,99}$ ways to arrange the students so that Aleksandra and Svetlana are in the first room and Alexey in the third room, etc. Thus the required probability is

$$\frac{2\binom{197}{48,50,99} + 2\binom{197}{48,49,100} + 2\binom{197}{49,50,98}}{\binom{200}{50,50,100}} \simeq 0.086$$

Notice that we could equivalently reasoned by arranging students in groups (not distinguishing the two groups of 50) instead of classrooms, which of course would have given us (of course it is the same)

$$\frac{\binom{197}{48,50,99} + \binom{197}{48,49,100} + \binom{197}{49,50,98}}{\binom{200}{50,50,100}/2}$$

P.7. To fix the notation, let's say that at the beginning there were n matches in each pocket. Let's calculate the probability that the person finds the *right* pocket empty, and that at this moment there are k matches in the left pocket. For this to happen, they must choose the right pocket n times out of $n+n-k$ tries, and then they must choose right yet again at the $2n-k+1$-th try. Thus this probability equals $\frac{1}{2}\binom{2n-k}{n}2^{-2n+k}$. The required probability is then $\binom{2n-k}{n}2^{-2n+k}$, since the left pocket may also be the first one to get empty.

P.8. If E is the set of passengers and F the set of wagons, the exercise is reduced to calculate some combinatorics of functions $f\colon E \to F$ (assignments of passengers into wagons). Since $|E| = k$ and $|F| = n$, there are k^n functions $f\colon E \to F$.

(8a) Here we need to calculate how many *surjective* functions $f\colon E \to F$ there are. Assume $k \geq n$ (otherwise the answer is 0). For $A \subset F$, define

$$\xi(A) = |\{\xi \in F^E \,:\, \xi(F) \subset A\}| = |A^E| = |A|^k$$
$$\eta(A) = |\{\xi \in F^E \,:\, \xi(F) = A\}|$$

Namely $\xi(A)$ is the number of functions with values in A, and $\eta(A)$ is the number of functions having exactly A as image. We are thus interested in $\eta(F)$.

Clearly $\xi(A) = \sum_{B\subset A} \eta(B)$. This implies from the inclusion-exclusion principle

$$\eta(F) = \sum_{A\subset F} (-1)^{|F|-|A|}\xi(A) = \sum_{j=0}^{n} (-1)^{n-j}\binom{n}{j} j^k \tag{5}$$

To find the probability, just divide by k^n.

(8b) We need to calculate how many functions have an image of cardinality r. From (5) and recalling that there are $\binom{n}{r}$ sets of cardinality r in F, we gather that there are

$$\sum_{B\subset F,\, |B|=r} \eta(B) = \sum_{j=0}^{r} (-1)^{n-j}\binom{n}{j, r-j, n-r} j^k$$

functions with an image of cardinality exactly r. This coincides with (8a) for $r = n$. Divide by k^n for the probability.

P.9. This is an exercise about the continuity of probability. We can think that the balls are numbered 1, 2 etc, (so the first second we introduce numbers from 1 to 100 etc).

(9a) Let E_n be the event where the ball number 1 is still in the urn after n seconds. There are $99 * 198 * \ldots = \prod_{i=1}^n 99$ n-extractions that leave the ball in the urn, out of the $\prod_{i=1}^n (1 + 99i)$ equally likely n-extractions. The probability is then

$$\mathbb{P}(E_n) = \prod_{i=1}^{n} \frac{99i}{1 + 99i}$$

(9b) Let A_i be the event where the ball number i stays in the urn forever. For instance the event A_i, corresponding to the ball number 1 staying in the urn forever, is $\cap_{n=1}^{\infty} E_n$, where E_n is the event in (9a). Since E_n is a decreasing sequence (and thus converging - recall the lecture 1 about *continuity of probability*) of events, we have

$$\mathbb{P}(A_1) = \mathbb{P}(\cap_{n=1}^{\infty} E_n) = \lim_n \mathbb{P}(E_n) = \lim_n \prod_{i=1}^{n} \frac{99i}{1 + 99i} \lim_n = \prod_{i=1}^{n} \frac{1}{1 + 1/(99i)} \leq \lim_n \frac{1}{\sum_{i=1}^{n} 1/(99i)} = 0$$

Similarly, if the ball m is introduced after k seconds, we have

$$\mathbb{P}(A_m) = \mathbb{P}(\cap_{n=k}^{\infty} E_n) \leq \lim_n \frac{1}{\sum_{i=k}^{n} 1/(99i)} = 0$$

So any given ball has probability 0 to stay in the urn forever. Beware, A_m is not the emptyset, it just has probability 0.

(9c) Notice that $A = \cup_i A_i$ is the event where at least one given ball stays in the urn forever. So $\mathbb{P}(A) \leq \sum_i \mathbb{P}(A_i) = 0$.

(9d) No! We are not saying that there are no balls in the urn.

P.10. In this problem it is easy to compute for a $B \subset F$, define $p_B := \sum_{y \in B} p_y$ and $A \subset E$

$$\mathbb{P}(f(A) \subset B) = (\sum_{y \in B} p_y)^{|A|} \tag{6}$$

However we rather want to know $\mathbb{P}(f(E) \supset B)$, and we will use inclusion-exclusion to express the latter in terms of (6).

Indeed

$$\mathbb{P}(f(A) \supset B) = \mathbb{P}(f(A) = B) + \mathbb{P}(f(A) \supset B, f(A) \neq B) \tag{7}$$

To compute the first term on the r.h.s. of (7)

$$\mathbb{P}(f(A) \subset B) = \sum_{B' \subset B} \mathbb{P}(f(A) = B')$$

so that by inclusion-exclusion

$$\mathbb{P}(f(A) = B) = \sum_{B' \subset B} (-1)^{|B| - |B'|} \mathbb{P}(f(A) \subset B') \tag{8}$$

To compute the second term in the r.h.s. of (7), we can identify $\bar{B}$ to a point denoted ζ with associated probability $p_\zeta = \sum_{t \in \bar{B}} p_b$. Then

$$\mathbb{P}(f(A) \supset B, f(A) \neq B) = \sum_{B' \subset B \cup \{\zeta\}} (-1)^{|B| + 1 - |B'|} \mathbb{P}(f(A) \in B')$$

In the last sum there are two type of terms: the B' that contain ζ, and those that do not contain it. The latter appear with the opposite sign that they have in (8). Therefore from (7)

$$\mathbb{P}(f(A) \supset B) = \sum_{B' \subset B} (-1)^{|B| - |B'|} \mathbb{P}(f(A) \in B' \cup \bar{B}) = \Big| \sum_{B' \subset B} (-1)^{|B'|} \mathbb{P}(f(A) \notin B') \Big|$$

So that

$$\mathbb{P}(f(E) \supset B) = \Big| \sum_{B' \subset B} (-1)^{|B'|} \big(\sum_{y \in \bar{B}'} p_y \big)^{|E|} \Big|$$

2 - Условная вероятность: Решения

P.11. This is $(1/6)/(1/2) = 1/3$ by symmetry.

P.12. They are trivially independent if $n < 7$. Otherwise, for $n \geq 7$, write $n = 14k+r$ with $0 \leq r < 14$; a direct computation shows that the events are independent if $r = 0, 2, 4, 6$.

P.13. First some useful notation. For integers $a \geq 0$, $b \geq -1$ and $c \geq 0$ denote

$$a^{(b,c)} := \prod_{i=0}^{c-1}(a + i\,b)$$

To make the notation clear, let us fix a precise probability space for this exercise. Let $\Omega := \{1, \dots, R\}^n$, an element $x = (x_1, \dots, x_n)$ of Ω just corresponds to the sequence of colors extracted. Since we take the discrete σ-algebra on Ω, it is enough to assign the probability of each singleton $\{x\}$ to specify the probability measure $\mathbb{P}$. Let then for a color $r \in \{1, \dots, R\}$

$$E_{i,r} := \{x \in \Omega \,:\, x_i = r\} = \text{‘the } i\text{-th ball is of color } r\text{’}$$

so that $\{x\} = \cap_{i=1}^n E_{i,x_i}$. The point here is that it is easy to calculate the probability to extract the color r at the i-th extraction, if we know what colors where extracted before: formally, this means conditioning. So we can use the following trivial identity

$$\mathbb{P}(\{x\}) = \mathbb{P}(\cap_{i=1}^n E_{i,x_i}) = \mathbb{P}(E_{1,x_1})\mathbb{P}(E_{2,x_2}|E_{1,x_1})\dots\mathbb{P}(E_{n,x_n}|\cap_{i=1}^{n-1} E_{i,x_i}) = \prod_{j=1}^n \mathbb{P}(E_{j,x_j} \cap_{i=1}^{j-1} E_{i,x_i}) \tag{9}$$

to make explicit calculations. Each of the probabilities on the r.h.s. is a rational number: the denominator of the j-th factor is just $((j-1)m + \sum_k u_k)$ since this is the number of balls in the urn at the j-th extraction. In the numerator, if the color i appears k times in the sequence x, then we have (somewhere) the factors $u_i(u_i+m)\dots(u_i+(k-1)m)$. So counting all the colors we get from (9)

$$\mathbb{P}(\{x\}) = \frac{\prod_{r=1}^R u_r^{(m,l_r(x))}}{u^{(m,n)}}$$

where $u = \sum_{i=1}^R u_i$ is the initial number of balls and

$$l_r \colon \Omega \to \{0, \dots, n\}$$
$$l_r(x) = \sum_{i=1}^n \mathbf{1}_{x_i = r} = \text{how many balls of color } r \text{ in } x$$

P.14. Let $p = \mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C)$. Then

$$(1-p)^2 = \mathbb{P}(\bar{B} \cap \bar{C}) \geq \mathbb{P}(\bar{A} \cap \bar{B} \cap \bar{C}) = 1 - \mathbb{P}(A \cup B \cup C) = 1 - (3p - 3p^2 + 0)$$

As a consequence, $p \leq 1/2$. It is easy to check that $p = 1/2$ is possible. E.g. $\Omega = \{1,2,3,4\}$ with uniform probability, and $A = \{1,2\}$, $B = \{1,3\}$, $C = \{2,3\}$.

P.15. Conditioning on the position of the first reshka, we have $p_n = p_{n-1}/2 + p_{n-2}/4 + p_{n-3}/8$ with $p_0 = p_1 = p_2 = 1$.

P.16. We chose randomly the head of each team, giving the same probability 1/15 to each member of each team. We then let A_i be the event “*the students at positions $i, i+1$ are not both heads of their groups*". We want to to show that $\mathbb{P}(\cap_{i=1}^{15000} A_i) > 0$. If the students at $i, i+1$ are in the same team, then $\mathbb{P}(\bar{A}_i) = 0$; otherwise $\mathbb{P}(\bar{A}_i) = 1/15^2$. In any case $\mathbb{P}(\bar{A}_i) \leq 1/15^2 =: p$. On the other hand, A_i is determined only by the choice of the heads in the teams of the students at positions $i, i+1$, and is independent on the choice of heads of the remaining 998 (or 999) teams. Since those teams are relevant for at most other $(2*(15-1)+1)*2 = 58$ adjacent positions $j, j+1$, we can take $d = 58$. Since p and d satisfy the conditions in Lovasz lemma, $\mathbb{P}(\cap_{i=1}^{15000} A_i) > 0$.

P.17. The key point in this exercise is not mathematical, but how we interpret the failure rate of the informant. We should understand that when the president is in the office (outside), the informants will report that he is outside (in the office) *indipendently* with probability 0.2 and 0.1. Formally, if we denote by A the event “*the president is in the office* and with A_1 (A_2) the event “*the first (second) informant* *reported that the president is in the office* we are assuming that $\mathbb{P}(A_1 \cap A_2|A) = \mathbb{P}(A_1|A)\mathbb{P}(A_2|A)$ etc. Then

$$\mathbb{P}(A|A_1 \cap \bar{A}_2) = \tfrac{\mathbb{P}(A_1\cap\bar{A}_2|A)\mathbb{P}(A)}{\mathbb{P}(A_1\cap\bar{A}_2|A)\mathbb{P}(A)+\mathbb{P}(A_1\cap\bar{A}_2|\bar{A})\mathbb{P}(\bar{A})}$$
$$= \tfrac{\mathbb{P}(A_1|A)\mathbb{P}(\bar{A}_2|A)\mathbb{P}(A)}{\mathbb{P}(A_1|A)\mathbb{P}(\bar{A}_2|A)\mathbb{P}(A)+\mathbb{P}(A_1|\bar{A})\mathbb{P}(\bar{A}_2|\bar{A})\mathbb{P}(\bar{A})} = \tfrac{0.8*0.1*0.6}{0.8*0.1*0.6+0.2*0.9*0.4}$$

P.18. Recall that the princess is only interested to choose to best candidate, namely she wants to maximize the probability to select the candidate ranked 1. The only thing she can choose is when to stop. It is clear that there is no point in choosing someone who is not the best ranked up to his arrival (as otherwise he is ranked 1 with probability 0). Due to permutation invariance, it is also clear that the optimal stopping strategy cannot depend on the (chronological) order of the ranks. Thus the optimal stopping strategy should be of the following type: *do not choose any of the first $M - 1$ candidates; starting from the M-th candidate, take the first one who betters all of those who came ahead of him. If none satisfies this constraint, take the last one.*

With this remark, one just needs to optimize over M. Let p_M be the probability to select the best using the aforementioned stopping policy depending on the parameter M. Denote B_i the event "*the i-th candidate is ranked* 1 and S_i the event "*the i-th candidate is selected*". Then

$$p_M = \sum_{i=1}^{N} \mathbb{P}(S_i|B_i)\mathbb{P}(B_i) = \sum_{i=M}^{N-1} \mathbb{P}(\text{the best of the first } i-1 \text{ is in the first } M-1)\tfrac{1}{N}$$
$$= \tfrac{1}{N} \sum_{i=M}^{N} \tfrac{M-1}{i-1}$$

(notice that the calculation is correct also for $i = N$). Thus the optimal M is the maximizer of $\frac{M-1}{N}\sum_{i=M-1}^{N-1}\frac{1}{i-1}$. For N large, it is easily seen that the optimal M satifies $\bar{M} = \frac{N}{e}(1 + o(1))$.

P.19. Let's discuss the point (19b). Let's call E_n the event "*after tossing a coin n times, we had Tail an even number of times* and let $r_n = \mathbb{P}(E_n)$. Then

$$\begin{aligned} r_0 &= 1 \\ r_{n+1} &= \mathbb{P}(E_{n+1}|E_n)\mathbb{P}(E_n) + \mathbb{P}(E_{n+1}|\bar{E}_n)\mathbb{P}(\bar{E}_n) \\ &= (1-p)r_n + p(1-r_n) \end{aligned}$$

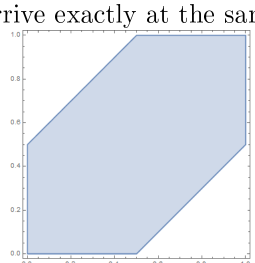
which implies $r_n = \frac{1+(1-2p)^n}{2}$.

P.20. Once the first 1 in the coin flip sequence appears, there are two possibilities for the next flip. If another 1 appears, then Bob wins for sure; if a 0 appears, then Bob still has some chances to win (if yet another 0 appears, Alice wins, but if 1 follows then the probability that Bob wins is the same as it was at the beginning). Thus Bob has the highest chance to win. More precisely, if p is the probability that Bob wins

$$p = \underbrace{\tfrac{1}{2}}_{\text{11 appears}} + \underbrace{\tfrac{1}{2}}_{\text{10 appears}} \Big(\underbrace{\tfrac{1}{2} * 0}_{\text{100 appears}} + \underbrace{\tfrac{1}{2}p}_{\text{101 appears}} \Big)$$

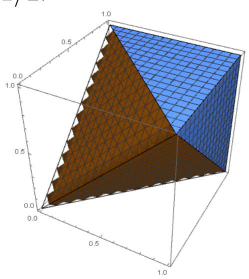
Thus $p = 2/3$.

P.21. The question is slightly ambiguous, as we do not know the probability distribution within the given time interval. In this cases, we interpret that it is somehow uniform (as when we say *pick a number between* 1 *and* 10!). It is then easily seen that the area of the shaded area in the picture is 3/4: it corresponds to the probability that X and Y meet each other. Similarly there is 0 probability that they arrive exactly at the same time.



P.22. Three numbers are the side lengths of a triangle iff the largest one does not exceed the sum of the other two. In other words, we have to calculate the volume of the region $\{\max\{x, y, z\} \le (x+y+z)/2\}$

within $[0,1]^3$. Since we are removing three tetahedrons, each of volume $1/6$, from such a cube, the answer is $1/2$.



3 - Случайные величины: Решения

P.23. Let N be the number of red balls drawn from the box in two extractions. We have $\mathbb{P}(N=1) = \frac{\binom{7}{1}\binom{5}{1}}{\binom{12}{2}} =: p_1$, and $\mathbb{P}(N=2) = \frac{\binom{7}{2}}{\binom{12}{2}} =: p_2$. This entails

$$\mathbb{M}[N] = p_1 * 1 + 2 * p2 = 7/6, \qquad \mathbb{D}[N] = p_1 * 1^2 + p_2 * 2^2 - \mathbb{M}[N]^2 = 175/396.$$

Notice that $N = N_1 + N_2$ where N_i takes the value 1 if the i-th ball was red, and 0 otherwise. Since $\mathbb{P}(N_i = 1) = 7/12 = 1 - \mathbb{P}(N_i = 0)$, from the linearity of the expectation it follows $\mathbb{M}[N] = \mathbb{M}[N_1] + \mathbb{M}[N_2] = 2 * 7/12$.

Now, if we sample with replacement (second question of the exercise), we still have $\mathbb{M}[N] = 7/6$ for the aforementioned reason. However, since now N_1 and N_2 are independent, we gater $\mathbb{D}[N] = \mathbb{D}[N_1] + \mathbb{D}[N_2] = 2 * (7/12)(1 - 7/12) = 35/72$.

Notice that in this second case the variance is higher, and this is quite intuitive. As a limiting case, you can think about sampling 12 balls. If there is no replacement, we will get exactly 7 red balls, thus the variance is 0. If we sample with replacement, then the variance grows with the number of extractions, as we get in this case $12 * (7/12(*1 - 7/12)$.

P.24. We have $\mathbb{M}[X+Y] = \mathbb{M}[X] + \mathbb{M}[Y] = \lambda_1 + 1/\lambda_2$ and by independence $\mathbb{D}(X+Y) = \mathbb{D}(X) + \mathbb{D}(Y) = \lambda_1 + \lambda_2^{-2}$. Similarly $\mathbb{M}[XY] = \mathbb{M}[X]\mathbb{M}[Y] = \lambda_1/\lambda_2$. $\mathbb{D}(XY) = \mathbb{M}[(XY)^2] - \mathbb{M}[XY]^2 = \mathbb{M}[X^2]\mathbb{M}[Y^2] - \mathbb{M}[XY]^2 = (\lambda_1 + \lambda_1^2) * 2\lambda_2^{-2} - \lambda_1^2\lambda_2^{-2}$.

Notice that $\mathbb{P}(\max(X,Y) \le x) = \mathbb{P}(X \le x)\mathbb{P}(Y \le x)$, therefore

$$\begin{aligned}\mathbb{M}[\max(X,Y)] &= \int_0^\infty \mathbb{P}(\max(X,Y) \ge x)dx = \int_0^\infty \mathbb{P}(X \ge x) + \mathbb{P}(Y \ge x)dx - \int_0^\infty \mathbb{P}(X \ge x)\mathbb{P}(Y \ge x)dx \\ &= \mathbb{M}[X] + \mathbb{M}[Y] - \int_0^\infty \mathbb{P}(X \ge x)e^{-\lambda_2 x}dx = \mathbb{M}[X] + \mathbb{M}[Y] - \mathbb{M}[\tfrac{1-e^{-\lambda_2 X}}{\lambda_2}] \\ &= \lambda_1 + e^{\lambda_1(e^{-\lambda_2}-1)}/\lambda_2\end{aligned}$$

Notice that this is larger than λ_1 and $1/\lambda_2$ as it should be, since $\mathbb{M}[\max(X,Y)] \ge \max(\mathbb{M}[X], \mathbb{M}[Y])$. A similar computation holds for the square.

$$\begin{aligned}\mathbb{M}[\max(X,Y)^2] &= \int_0^\infty \mathbb{P}(\max(X^2,Y^2) \ge x)\,dx \\ &= \mathbb{M}[X^2] + \mathbb{M}[Y^2] - \int_0^\infty \mathbb{P}(X^2 \ge x)e^{-\lambda_2\sqrt{x}}dx = \mathbb{M}[X^2] + \mathbb{M}[Y^2] - \tfrac{2}{\lambda_2^2}(1 - \mathbb{M}[e^{-\lambda_2 X}(1+\lambda_2 X)] \\ &= \lambda_1 + \lambda_1^2 + \tfrac{2e^{\lambda_1(e^{-\lambda_2}-1)}}{\lambda_2^2} + \tfrac{2\lambda_1}{\lambda_2}\exp(e^{-\lambda_2}\lambda_1 - \lambda_2 - \lambda_1)\end{aligned}$$

P.25. Let $Y_1, \dots Y_n \in \{0,1\}$ be the results of the students in the first class (1 means they passed the exam), and let $X_1, \dots, X_N$ the the results for the second class. Then $Y = \sum^n Y_i$ and $X = \sum^N X_i$ are the numbers of students that passed the exams. Then by linearity $\mathbb{M}[Y] = n\,p$ and by independence $\mathbb{D}(Y) = n\,p(1-p)$.On the other hand

$$\mathbb{M}[X] = \mathbb{M}[\mathbb{M}[X|N]] = \sum_{k=0}^\infty \mathbb{M}[X|N=k]\mathbb{P}(N=k) = \sum_{k=0}^\infty k\,p\,\mathbb{P}(N=k) = \mathbb{M}[N]\,p = n\,p = \mathbb{M}[Y]$$

Similarly

$$\begin{aligned}\mathbb{M}[X^2] &= \sum_{k=0}^\infty \mathbb{M}[X^2|N=k]\mathbb{P}(N=k) \\ &= \sum_{k=0}^\infty (k\,p + k(k-1)p^2)\mathbb{P}(N=k) = \mathbb{M}[N]\,p(1-p) + \mathbb{M}[N^2]p^2 = \mathrm{Var}[Y] + \mathbb{M}[N^2]p^2\end{aligned}$$

Therefore $\mathrm{Var}[X] = \mathrm{Var}[Y] + p^2\mathrm{Var}[N] \ge \mathrm{Var}[Y]$.

P.26. If we select a permutation over n elements (here $n = 100$) randomly (more precisely, giving to each permutation the same probability $1/n!$), the probability that a given number $i \in \{1, \dots, n\}$ is fixed by the permutation is $1/n$. Again, by linearity of the expectation, the expected number of fixed points is $n\,\frac{1}{n} = 1$, regardless of n.

As for the variance, let X_i take the value 1 (if i is a fixed point) or 0 (othervise). We have $\mathbb{M}[X_i^2] = \mathbb{M}[X_i] = 1/n$. On the other hand, for $i \neq j$, $\mathbb{M}[X_i\, X_j] = \frac{1}{n(n-1)}$ (since it coincides with the probability that i and j are both fixed). Therefore the variance of the number of fixed points $X = \sum_i X_i$ is calculated as

$$\mathbb{D}[X] = \sum_{i=1}^{n} \mathbb{M}[X_i^2] + \sum_{i \neq j}^{n} \mathbb{M}[X_i\, X_j] \;- \mathbb{M}[X]^2$$
$$n\tfrac{1}{n} + n(n-1)\tfrac{1}{n(n-1)} - 1 = 1$$

By inclusion-exclusion, the probability that a permutation over N elements has no fixed point is $\sum_{k=0^N} \frac{(-1)^k}{k!}$. It follows that the probability that a permutation over n elements has exactly ℓ fixed points is

$$p_\ell = \tfrac{1}{n!} \underbrace{\binom{n}{\ell}}_{\text{choose } \ell \text{ fixed points}} \underbrace{(n-\ell)! \sum_{k=1}^{n-\ell} \tfrac{(-1)^k}{k!}}_{\text{perm. with no fixed points on } (n-\ell) \text{ el.}} = \tfrac{1}{\ell!} \sum_{k=1}^{n-\ell} \tfrac{(-1)^k}{k!}$$

We have thus gathered for $n \geq 2$

$$\mathbb{M}[1] = \sum_\ell p_\ell = 1$$
$$\mathbb{M}[X] = \sum_\ell p_\ell\, \ell = 1$$
$$\mathbb{M}[X^2] = \sum_\ell p_\ell\, \ell^2 = 2$$

Can you compute $\mathbb{M}[X^k]$ for $n \geq k$? [Hint: expand the power as we did for the square, to find the Bell number B_k]

P.27. If X is the number of tosses, then we have for (a)

$$p_n := \mathbb{P}_N(X = n) = \begin{cases} (1-p)^{n-1}p & \text{if } 1 \leq n \leq N-1 \\ (1-p)^{N-1} & \text{if } n = N \end{cases} \tag{10}$$

Thus we have

$$\begin{aligned} \mathbb{M}_N[X] &= \sum_{n=1}^{N} np_n = \tfrac{1-(1-p)^N}{p} \\ \mathbb{D}_N[X] &= \sum_{n=1}^{N} n^2 p_n - \mathbb{M}[X]^2 = \frac{1-p-p(1-p)^N(2N-1)-(1-p)^{2N}}{p^2} \end{aligned} \tag{11}$$

The case (b) is similar, just we get series at the place of finite sums. In this case $\mathbb{M}[X] = 1/p$ and $\mathbb{D}[X] = (1-p)p^{-2}$.

P.28. It is a continuous, increasing function. Beware, it is continuous even if Poisson is discrete.

P.29. If we pick $X_1, \ldots, X_n$ independent with uniform distribution, then

$$\begin{aligned} \mathbb{M}[|\cup_{i=1}^n (X_i + E)|] &= \mathbb{M}[\int \mathbf{1}_{\cup_i (X_i+E)}(x)dx] = \int \mathbb{M}[\mathbf{1}_{\cup_i (X_i+E)}(x)]dx \\ &= \int 1 - \mathbb{M}[\mathbf{1}_{\cap_i (X_i+E)^c}(x)]dx = 1-(1-|E|)^n \end{aligned}$$

So there is at least one point where the inequality holds. Exactly the same proof holds for compact groups with a Haar measure. If we take $|E| = 1/n$, we then get an elementary proof of a result due to J.Bourgain.

P.30. If one assume that X admits a density, one can make the computation explicitly by taking derivative (which is also a good exercise). Let us solve it however in the general case.

(30a) Take $a \leq b$. Then

$$\begin{aligned} \mathbb{M}[|X-a|] - \mathbb{M}[|X-b|] &= \mathbb{M}[(a-b)\mathbf{1}_{X\leq a}] + \mathbb{M}[(b-a)\mathbf{1}_{X \geq b}] + \mathbb{M}[(2X-a-b)\mathbf{1}_{a<X<b}] \\ &\geq (b-a)(\mathbb{P}(X \geq b) - \mathbb{P}(X \leq a) - \mathbb{P}(a < X < b)) = (b-a)(\mathbb{P}(X \geq b) - \mathbb{P}(X < b)) \end{aligned}$$

If b is larger than any median a, the last expression is negative. So $\mathbb{M}[|X-m|] \leq \mathbb{M}[|X-b|]$ for b larger than m. Similarly for b smaller (or applying the same result to $-X$). Computing the derivative is more or less acceptable (to be precise, one must notice that ϕ_1 is convex and compute a subdifferential - which is what some people did, using different words).

(30b) We have seen it many times. Just develop the square $\mathbb{M}[|X - \mathbb{M}[X] + \mathbb{M}[X] - r|^2]$.

(30c) For (30a), $|m - \mathbb{M}[X]|^2 \le \mathbb{M}[|X - m|]^2 \le \mathbb{M}[|X - \mathbb{M}[X]|]^2 \le \mathbb{D}(X)$

P.31. Integrating in $\mathbb{R}^d$ by parts, we see that the formula

$$\mathbb{M}[\operatorname{div}(Sf(X))] = \mathbb{M}[X \cdot f(X)]$$

holds for $f\colon \mathbb{R}^d \to \mathbb{R}^d$, and any (invertible or not) S.

P.32. For n odd we have 0. The homogeneity in σ is immediate, for instance considering $Y = X/\sigma \sim \mathcal{N}(0,1)$. Then we have $\mathbb{M}[f''(Y)] = \mathbb{M}[Yf'(Y)] = \mathbb{M}[Y^2 f(Y)]$, so that by induction for $f(Y) = Y^n$ we get that the answer is $\sigma^n(n-1)!!$.

P.33. Let Y be an independent copy of X. Then $(f(X)-f(Y))(g(X)-g(Y)) \ge 0$, since f is monotone. Therefore $0 \le \mathbb{M}[(f(X) - f(Y))(g(X) - g(Y))] = 2\mathbb{M}[f(X)g(X)] - 2\mathbb{M}[f(X)]\mathbb{M}[g(X)]$.

P.34. We can take $\mathbb{M}[X] = 0$. It is clear that it is convenient to concentrate all the mass at 0 and at the points $\pm c$. So we take $\mathbb{P}(X = +c) = \mathbb{P}(X = -c) = c^2/(2a)$ and $\mathbb{P}(X = 0) = 1 - c^2/a$.

P.35. This is a classical problem, as it has been used to provide the first Monte Carlo simulation in history. Here we discuss a slight variation of this problem: suppose that the '*needle*"is replaced by any Lipschitz curve of length ℓ. What is the expected number of intersections with a grid of horizontal lines, spaced at distance 1 from each other?

The answer is surprisingly simple. First, if the needle is just a segment of length ℓ, the expected value must be proportional to ℓ, $\mathbb{M}[N] = c\ell$ for some $c \ge 0$ (it is an increasing function of ℓ and satsfies $f(\ell_1+\ell_2) = f(\ell_1)+f(\ell_2)$). Next if the needle is a polygonal curve of total length ℓ, then again by linearity of the expected value $\mathbb{M}[N] = c\ell$ since the total number of intersections is the sum of the intersections of each segment in the needle. Finally, by monotone approximation this holds for any sufficiently regular curve. Thus the whole problem is reduced to calculate the constant c. However, if the needle is a circle of radius $1/2$, then $N = 2$ a.s., and therefore in this case $2 = \mathbb{M}[N] = c\,(2\pi)$. So that $c = 1/(2\pi)$, and the expected value of the number of intersections for any Lipschitz curve is $\mathbb{M}[N] = \frac{\ell}{2\pi}$. Notice that if the needle is a segment of length $\ell < 1$, then this also coincides with the probability of having at least one intersections (since two or more intersections cannot occur in this case).

P.36. Let X be a random variable with values in a measurable space E, and let $\mathbb{P}$ be the law of X (thus $\mathbb{P}$ is a probability measure on E). If a map $f\colon E \to F$ is measurable, then the law of the random variable $f(X)$ is $\mathbb{P} \circ f^{-1}$. In particular, let's consider the case $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$ (the same statements hold on manifolds), let's assume that $\mathbb{P}$ admits a density φ w.r.t. the Lebesgue measure, that the map f is smooth a.e. and (locally) injective. Then also the law of the random variable $U = f(X)$ admits a density ψ on $\mathbb{R}^m$, and the usual change of variables formula holds

$$\psi(y) = \frac{\varphi \circ f^{-1}(y)}{|Df| \circ f^{-1}(y)}, \tag{12}$$

where $|Df|$ stands for the absolute value of the Jacobian determinant of f. The exercise is the solved using $\varphi(x,y) = \mathbf{1}_{0\le x,y\le 1}$, and f the various maps given in the exercise.

P.37. We may apply directly the computation (12). To avoid cumbersome computations, let's split the work in two parts. First, define the random variable $A = \sqrt{-2\log\xi}$. Notice that for $a \ge 0$

$$\mathbb{P}(A \le a) = \mathbb{P}(\xi \ge e^{-a^2/1}) = 1 - e^{-a^2/2}.$$

In particular the law of A admits a density $\chi(a) = a\,e^{-a^2/2}\mathbf{1}_{a\ge 0}$. Since ξ and η are independent, also A and η are. Therefore (A,η) admits the density

$$\varphi(a,s) = a\,e^{-a^2/2}\mathbf{1}_{a\ge 0}\mathbf{1}_{0\le s\le 1}.$$

Now let's apply (12) to the transformation

$$f(a,s) = (a\,\cos(2\pi s), a\,\sin(2\pi s))$$

which is just the usual polar coordinates transformation. An immediate application of (12) then gives for (X,Y) the standard Gaussian density

$$\psi(x,y) = \frac{1}{2\pi}\exp(-\tfrac{x^2+y^2}{2}).$$

P.38. No for instance

$$\mathbb{P}(X > 2^{-1/2}) = \mathbb{P}(Y > 2^{-1/2}) > 0$$
$$\mathbb{P}(X > 2^{-1/2}, Y > 2^{-1/2}) = 0$$

P.39. Recall that the density of Y_n is defined, if it exists, by the property $\mathbb{P}(Y_n \in [a,b]) = \int_a^b \varrho_n(y)\,dy$. Otherwise said $\varrho_n(y) = -\frac{d}{dy}F_{Y_n}(y)$ provided the distribution of Y_n admits a density.

(39a) By induction: for $n = 1$, this is just the density of an exponential random variable. Let us assume the result for n, and notice that $\frac{de^{\lambda y}\varrho_{n+1}(y)}{dy} = e^{\lambda y}\varrho_n(y)/\lambda$ for $y > 0$. Then we have for $n+1$

$$\mathbb{P}(Y_{n+1} \le t) = \mathbb{P}(Y_n + X_{n+1} \le t) = \int_{x,y\ge 0, y+x\le t} \underbrace{\varrho_n(y)\varrho_1(x)}_{independence}\,dy\,dx$$
$$= \int_0^t dy\, e^{-\lambda y}\tfrac{d}{dy}(e^{\lambda y}\varrho_{n+1}(y))\int_0^{t-y}\lambda e^{-\lambda x}dx = \int_0^t dy\,\varrho_{n+1}(y)$$

where we integrated by parts in the last equality.

(39b) Let $E_{n,t} := \{Y_n > t\}$. Then $\mathbb{P}(N_t = n) = \mathbb{P}(E_{n+1,t} \cap E^c_{n,t}) = \mathbb{P}(E_{n+1,t}) - \mathbb{P}(E_{n,t})$.

However from (39a) we have that

$$\mathbb{P}(E_{n+1,t}) = e^{-\lambda t}\sum_{k=0}^{n}(\lambda t)^k/k!$$

as can be seen by deriving w.r.t. t. So the result follows.

P.40. Notice that $\mathbb{M}[e^{rX}] < \infty$ only entails a bound for the positive part of X.

(40a) By the Cauchy-Schwarz inequality

$$e^{\psi((1-\alpha)t+\alpha s)} = \mathbb{M}\big[e^{((1-\alpha)t+\alpha s)X}\big]$$
$$\le \mathbb{M}\big[(e^{(1-\alpha)tX})^{\frac{1}{1-\alpha}}\big]^{(1-\alpha)}\mathbb{M}\big[(e^{\alpha s X})^{\frac{1}{\alpha}}\big]^{\alpha} = e^{(1-\alpha)\psi(t)+\alpha\psi(s)}$$

(40b) Since $\psi(0) = 0$, one gets the inequality $\psi^* \ge 0$ taking $t = 0$ in the supremum in (3). On the other hand, by Jensen inequality

$$\psi(t) = \log\mathbb{M}[e^{tX}] \ge \log e^{t\mathbb{M}[X]} = t\mathbb{M}[X]$$

This means that for $x \le \mathbb{M}[X]$, one has $t\,x - \psi(t) \le 0$. Therefore, for such an x, the supremum in (3) is achieved for $t = 0$.

(40c) For all $t \in [0, r)$ it holds

$$\mathbb{P}(X \ge x) \le \mathbb{P}(e^{tX} \ge e^{tx}) \le e^{-tx}\,\mathbb{M}[e^{tX}] = e^{-(tx-\psi(t))}$$

Optimize over t (namely take the inf in the r.h.s.), to get the result. It is then easy to check that $\psi_N(t) := \log\mathbb{M}[e^{t(X_1+\ldots+X_N)/N}] = N\psi(t/N)$, so that $\psi^*_N(t) = N\psi^*(t)$.

(40d) Assume $\bar{r} > 0$ or there is nothing to prove. Since $|x|^n \le r^{-n}\,n!\,e^{r|x|}$, we have that $\mathbb{M}[|X|^n]$ is finite, and for any $r \in (0, \bar{r})$ one has

$$m_n := \mathbb{M}[X^n] \le r^{-n}\,n!\,\mathbb{M}[e^{r|X|}]$$

Therefore the power series

$$\phi(z) := \sum_{k=0}^{\infty}\tfrac{m_n}{n!}z^n$$

is dominated by a geometric series and converges uniformly in any ball $\{|z| \le R\}$ of radius $R < \bar{r}$. Thus ϕ is analytic in $\{|z| < \bar{r}\}$. By dominated and monotone convergence, ϕ coincides with ψ on the real segment $(0, \bar{r})$.

(40e) For the first example, consider an exponential random variable of parameter $\lambda > 0$. One has $\bar{r} = \lambda$ and $\mathbb{M}[e^{tX}] = \int_{\mathbb{R}^+}\lambda e^{(t-\lambda)x}dx = \frac{\lambda}{\lambda - t} \uparrow +\infty$ as $t \uparrow \lambda$.

For the second example, consider a random variable with density $\varrho(x) = \frac{1}{Z}\frac{e^{-\lambda x}}{1+x^2}\mathbf{1}_{x\ge 0}$, where Z is a suitable constant. Then $\bar{r} = \lambda$ but for $t < \lambda$

$$\mathbb{M}[e^{tX}] \le \int_{\mathbb{R}^+}\tfrac{1}{Z(1+x^2)}dx < \infty$$

so that $\lim_{t\uparrow\bar{r}} \psi(t)$ exists (ψ is increasing in this case) and is finite.

P.41. Let us clear out the general situation. Let U be a random variable and $V = f(U)$ for some measurable function f. Then U are V are independent iff V is a.s. constant. Indeed, if V is constant U and V are trivially independent. Conversely, if U and V are independent, then for every measurable set A

$$\begin{aligned} 0 &= \mathbb{P}(V \in \bar{A}, V \in A) = \mathbb{P}(U \in f^{-1}(\bar{A}), V \in A) \\ &= \mathbb{P}(U \in f^{-1}(\bar{A}))\mathbb{P}(V \in A) = \mathbb{P}(V \in \bar{A})\mathbb{P}(V \in A) \end{aligned}$$

Thus for every measurable A the probability of the event $\{V \in A\}$ is either 0 or 1, which easily implies that V is constant a.s..

Notice that the same statement does not hold in the following situation. U is a random variable, and we know that $V = g(U)$ and $W = f(U)$ are independent. In this case we *cannot* deduce that either V or W are a.s. constant. For instance take U uniform in $[-1, 1]$, $V = \mathrm{sign}(U)$ and $W = |V|$.

Back to exercise, if X and Y are independent, then X^2 and Y^2 are also independent. And if $X^2+Y^2 = 1$, then from the previous remark both X^2 and Y^2 are constant a.s., say $X^2 = \cos^2(\theta)$ and $Y^2 = \sin^2(\theta)$ for some angle θ. Thus it must happen $X = U\cos(\theta)$ and $Y = V\sin(\theta)$ where U, V are independent random variables with (a.s.) values in $\{-1, +1\}$.

P.42. Reasoning as in the previous exercise, it follows that X a.s. takes value on the inverse image of sin of some given value $u \in [-1, 1]$. We can describe this as follows. There exists a (deterministic) angle $\theta \in S^1$, and a Bernoulli random variable U and an integer valued random variable K such that $X = 2\pi K + U\theta + (1 - U)(\pi/2 - \theta)$.

P.43. The energy is the random variable $e = \int_0^T w_t\,dt$. So

$$\begin{aligned} \mathbb{M}[e] &= \mathbb{M}[\mathbb{M}[e|T]] = \mathbb{M}[\mathbb{M}[\int_0^T w_t\,dt|T]] = \mathbb{M}[\int_0^T \mathbb{M}[w_t|T]\,dt] \\ &= \mathbb{M}[\int_0^T \mathbb{M}[w_t\,dt]] = \mathbb{M}[\frac{T(a+b)}{2}] = \frac{a+b}{2\lambda} = 4 * 10^7\,s\,w \end{aligned}$$

4 - Характеристические функции: Решения

P.44. Let ξ, η be two independent standard normal random variables, and let U and V be two random variables such that $U^2 + V^2 = 1$ a.s., and assume that (U, V) and ξ and η are independent (beware, we are not assuming that U and V are independent; we are assuming that the σ-algebras generated by (U, V), η and ξ are independent). We will prove that $\xi U + \eta V$ is a standard gaussian random variable. Then taking $U = \cos(\alpha)$ and $V = \sin(\alpha)$ we will have solved the exercise.

Indeed, let $\mathcal{G}$ be the σ-algebra generated by (U, V)

$$\begin{aligned}\varphi_{\xi U+\eta V}(t) &= \mathbb{M}\Big[\mathbb{M}\big[e^{it\xi U}e^{it\eta V}|\mathcal{G}\big]\Big] = \mathbb{M}\Big[\mathbb{M}\big[e^{it\xi U}|\mathcal{G}\big]\mathbb{M}\big[e^{it\eta V}|\mathcal{G}\big]\Big] \\ &= \mathbb{M}\big[e^{-\frac{t^2U^2}{2}}e^{-\frac{t^2V^2}{2}}\big] = e^{-\frac{t^2}{2}}\end{aligned}$$

where in the first line we used that ξU and ηV are independent conditionally to $\mathcal{G}$. Since the characteristic function of $\xi U + \eta V$ coincides with the standard normal one, $\xi U + \eta V$ is a standard normal variable.

If you find the notation in (10) confusing, we can write down things explicitly (although the probabilistic proof (10) is slightly more general then the one below). Let γ be the standard Gaussian measure ($d\gamma(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}}dx$), and ν be the law of (U, V); we are assuming in particular

$$\nu(\{(u, v)\,:\,u^2 + v^2 = 1\}) = 1$$

(in other words, ν is concentrated on a circle). Then

$$\begin{aligned}\varphi_{\xi U+\eta V}(t) &= \int d\nu(u,v)\,d\gamma(x)\,d\gamma(y)e^{it\,(x\,u+y\,v)} = \int d\nu(u,v)\Big[\Big(\int d\gamma(x)e^{i\,t\,u\,x}\Big)\Big(\int d\gamma(y)e^{i\,t\,v\,y}\Big)\Big] \\ &= \int d\nu(u,v)e^{-t^2u^2/2}e^{-t^2v^2/2} = e^{-t^2/2}\end{aligned}$$

If you still fill confused about the notation, let's write what (10)-(11) mean. Let ϱ be the gaussian density function. Then

$$\begin{aligned}\varphi_{\xi\cos(\alpha)+\eta\sin(\alpha)}(t) &= \int_{\mathbb{R}\times\mathbb{R}\times S^1} dx\,dy\,d\theta\,\varrho(x)\varrho(y)\tfrac{1}{2\pi}e^{it(x\cos(\theta)+y\sin(\theta))} \\ &= \int_{S^1} d\theta\,\tfrac{1}{2\pi}\Big[\int_{\mathbb{R}} dx\varrho(x)e^{itx\cos(\theta)}\int_{\mathbb{R}} dx\varrho(y)e^{ity\sin(\theta)} \\ &= \int_{S^1} d\theta\,\tfrac{1}{2\pi}\Big[e^{-\frac{t^2\cos^2(\theta)}{2}}e^{-\frac{t^2\sin^2(\theta)}{2}}\Big] = e^{-\frac{t^2}{2}}\end{aligned}$$

P.45. The same as before, taking $U = 1/\sqrt{1+Z^2}$ and $V = Z/\sqrt{1+Z^2}$.

P.46. By induction, it is enough to do it for two random variables. Let X_1 and X_2 be independent Poisson with parameters λ_1 and λ_2 respectively. Then from the independence of X_1 and X_2 and **Ex 1-(c)**

$$\begin{aligned}\varphi_{X_1+X_2}(t) :&= \mathbb{M}[e^{it(X_1+X_2)}] = \mathbb{M}[e^{itX_1}]\mathbb{M}[e^{itX_2}] \\ &= \exp((e^{it}-1)\lambda_1)\exp((e^{it}-1)\lambda_2) = \exp((e^{it}-1)\end{aligned}$$

which is indeed the characteristic function of a Poisson variable of intensity $\lambda_1 + \lambda_2$.

As for the counterexample, just take $X_1 = X_2$. The sum is even with probability 1.

P.47. $\sin(0) = 0 \neq 1$, $\sin(t^2)$ is not uniformly continuous, $\sup_t |1+\sin(t)| = 2 > 1$, $\frac{d}{dt}e^{-t^4}|_{t=0} = -24 < 0$ cannot be 4! a fourth moment.

P.48. We can take $X \sim \mathcal{N}(0,1)$ and U a Bernoulli random variable of parameter $1/2$ independent of X. Then $Y = (2U-1)X$. Then

$$\mathbb{M}[e^{itY}] = \mathbb{M}[e^{it(2U-1)X}] = \tfrac{1}{2}\mathbb{M}[e^{itX}] + \tfrac{1}{2}\mathbb{M}[e^{-itX}] = e^{-t^2/2}$$

So $X, Y \sim \mathcal{N}(0,1)$. Moreover $\mathbb{M}[XY] = \mathbb{M}[2U-1]\mathbb{M}[X^2] = 0$. But

$$\mathbb{M}[e^{isX+itY}] = \mathbb{M}[e^{i(s+(2U-1)t)X}] = \tfrac{1}{2}\mathbb{M}[e^{i(s-t)X}|U=0] + \tfrac{1}{2}\mathbb{M}[e^{i(s+t)X}|U=1] = e^{-(t^2+s^2)/2}\cosh(ts)$$

which is not the exponential of a quadratic form.

P.49. Let $f(z) := \frac{e^{itz}}{\pi(1+z^2)}$. f has simple poles at $z = \pm i$, we are interested at the residue at $z = i$, namely

$$\mathrm{Res}_{z=i}f(z) = \lim_{z\to i}(z-i)f(z) = \tfrac{e^{-t}}{2\pi i}$$

Let $A_r := \{z \in \mathbb{C}, : \operatorname{Im}(z) \geq 0, |z| \leq r\}$, and let C_r be the boundary of A_r with anticlockwise orientation. Then

$$e^{-t} = 2\pi i \operatorname{Res}_{z=i} f(z) = \int_{C_r} f(z)dz = \int_{-r}^{r} \tfrac{e^{itx}}{\pi(1+x^2)} dx + \int_{C_r \setminus [-r,r]} f(z)dz$$

If $t > 0$, it is easy to see that the last integral vanishes as $r \to \infty$ (since $t > 0$ and $\operatorname{Im}(z) \geq 0$, $|e^{itz}| \leq 1$; then we integrate a function decaying as r^{-2} along a contour of length $\sim r$); while the first integral converges to $\varphi_X(t)$ as $r \to \infty$. Since it is immediate to check that φ_X is even (as the density is even), we gather $\varphi_X(t) = \exp(-|t|)$. The lack of differentiability is due to the fact that $\mathbb{M}[|X|] = +\infty$.

P.50. Notice that the point X where the cannonball falls has Cauchy distribution since $X = \tan(\theta)$.

(50a) The value of the damage is V_X. Therefore

$$\mathbb{M}[V_X] = \mathbb{M}[\mathbb{M}[V_X|X]] = \mathbb{M}[1/\lambda_X] = \int \frac{1}{\pi(1+x^2)}(1+x^2)\exp(-|x|)dx = 2/\pi$$

Similarly

$$\mathbb{M}[V_X^2] = \mathbb{M}[\mathbb{M}[V_X^2|X]] = \mathbb{M}[2/\lambda_X^2] = \int \frac{2}{\pi(1+x^2)}(1+x^2)^2\exp(-2|x|)dx = 3/\pi$$

So $\mathbb{D}(V_X) = 1/\pi$.

(50b) Let us call X and Y the points where the first and second cannonball fall. The total damage is $V_X + V_Y$. By linearity $\mathbb{M}[V_X + V_Y] = 2\mathbb{M}[V_X] = 4/\pi$.

$$\mathbb{D}[V_X + V_Y] = \mathbb{M}[V_X^2] + \mathbb{M}[V_Y^2] + 2\mathbb{M}[V_X V_Y] - \mathbb{M}[V_X + V_Y]^2 = 2\mathbb{D}[V_X] + 2(\mathbb{M}[V_X V_Y] - \mathbb{M}[V_X]\mathbb{M}[V_Y])$$

So we only need to calculate

$$\mathbb{M}[V_X V_Y] - \mathbb{M}[V_X]\mathbb{M}[V_Y] = \mathbb{M}[\mathbb{M}[V_X V_Y - \mathbb{M}[V_X]\mathbb{M}[V_Y]|(X,Y)]] = \mathbb{M}[\frac{e^{-|X-Y|}}{\lambda_X \lambda_Y}]$$
$$= \int \pi^{-2}\exp(-|x| - |y| - |x-y|)dxdy = 3/(4\pi^2)$$

5 - Условное мат.ожидание: Решения

P.51. By the definition of conditional expectation, we need to find a measurable map (up to a.e. equivalence) $q\colon F \to \mathbb{R}$ such that

$$\int_{E\times F} g(y)f(x,y)d\theta(x,y) = \int_F g(y)q(y)d\theta(E,y)$$

for all bounded measurable g. It follows

$$\mathbb{M}[f(X,Y)|Y](y) = q(y) = \frac{\int_E f(x,y)\varrho(x,y)d\mu(x)}{\int_E \varrho(x,y)d\mu(x)}$$

P.52. Yes, it is Gaussian. We can use the previous exercise to prove it, with $\mu = \nu =$Lebesgue. Then the conditional distribution will have density

$$\varrho(x,Y)/\int \varrho(x',Y)dx'$$

However, instead of solving the integral, let us reason as follows. In the previous integral the quadratic term in front of x is independent of Y. So, if X is still Gaussian conditionally to Y, the variance of the conditional distribution is independent of Y. So if we let $s_{11} = \mathbb{D}(X)$, we try to check if a $\mathcal{N}(aY+b, s_{11})$ may be the conditional distribution of X given Y, for some $a, b \in \mathbb{R}$. In this case we would have

$$\mathbb{M}[e^{itX}|Y] = e^{it(aY+b)-\frac{1}{2}s_{11}t^2}$$

So we only need to check if there exist a and b, such that for all $t, u \in \mathbb{R}$

$$\mathbb{M}[e^{itX}e^{iuY}] = \mathbb{M}[e^{it(aY+b)-\frac{1}{2}s_{11}t^2}e^{iuY}]$$

With obvious notation the left hand side is

$$\log \mathbb{M}[e^{itX}e^{iuY}] = it\,m_1 + iu\,m_2 - \tfrac{1}{2}s_{11}t^2 - \tfrac{1}{2}s_{22}u^2 - s_{12}tu$$

The right hand side

$$\log \mathbb{M}[e^{i(ta+u)Y+itb-\frac{1}{2}s_{11}t^2}] = itb - \tfrac{1}{2}s_{11}t^2 + i(ta+u)m_2 - \tfrac{1}{2}s_{22}(ta+u)^2$$

The two are equals for $a = s_{12}/s_{22}$ and $b = m_1 - am_2$. So X conditionally to $Y = y$ is Gaussian with mean $\mathbb{M}[X] + \frac{\mathrm{Cov}(X,Y)}{\mathbb{D}(Y)}(y - \mathbb{M}[Y])$ and variance $\mathbb{D}(X)$. The correction term means 0 if $\mathbb{D}(Y) = 0$.

P.53. We have that $H_a = X + Y_a$ and $H_b = X + Y_b$ with X, Y_a, Y_b independent. We are asked to characterize $\mathbb{P}(H_b \in \cdot|H_a)$.

Notice that (H_a, H_b) is obtained from the Gaussian vector (X, Y_a, Y_b) via a linear transformation. It is therefore Gaussian with mean $(2m, 2m)$ and covariance S with entries $S_{i,j} = \mathbb{M}[(H_i - m)(H_j - m)]$. So $S_{a,a} = \mathbb{D}(X + Y_a) = \mathbb{D}(X) + \mathbb{D}(Y_a) = 2\sigma^2$. Similarly $S_{b,b} = 2\sigma^2$, while $S_{a,b} = S_{b,a} = \mathbb{M}[(X + Y_a - 2m)(X + Y_b - 2m)] = \sigma^2$.

We can then use the previous exercise, to find that X is Gaussian with mean $2m + \frac{S_{ab}}{S_{bb}}(y - 2m) = 2m + (y - 2m)/2 = 175$cm.

As an alternative, if we do not want to use 'Gaussianity', we could have found the same with a more general procedure. Notice that the transformation $(x, y_a, y_b) \to (x, x + y_a, x + y_b)$ has determinant 1. Therefore, for v the Gaussian density in the text, the law of (H_a, H_b) admits the density $\varrho(h_a, h_b) = \int v(x)v(x - h_a)v(x - h_b)dx$, which coincides with the Gaussian density with mean $(2m, 2m)$ and variance S.

Therefore, as in the previous exercises, we have

$$\mathbb{P}(H_a \in A|H_b) = \mathbb{M}[\mathbf{1}_A(H_a)|H_b] = \frac{\int_A \varrho(h, H_b)dh}{\int \varrho(h', H_b)dh'}$$

Thus

$$\mathbb{M}[H_a|H_b] = \tfrac{\int h\varrho(h,H_b)dh}{\int \varrho(h,H_b)dh}$$

P.54. We claim $\bar{Y} = \mathbb{M}[X|\mathfrak{G}]$. Indeed, write $Y = \bar{Y} + Z$ to see that $\mathbb{M}[(X - Y)^2] = \mathbb{M}[(X - Y)^2 + Z^2] + \mathbb{M}[\mathbb{M}[Z(X - \bar{Y})|\mathfrak{G}]]$, and the last term vanishes.

6 - Сходимость: Решения

P.55. It happens if $a < 0$ and $b > 0$. Indeed, it is easily seen that if $a < 0$ and $b > 0$, then $\int_a^b \exp(-nx^2/2)dx/Z_n \to 1$. If $a, b > 0$ or $a, b < 0$ then the limit is 0, as it is for $a = b$. Finally for $a = 0$ and $b > 0$ or $a < 0$ and $b = 0$, the limit is 1/2. On the other $\mathbb{P}(X \in [a,b]) = \delta_0([a,b])$. So we need $a \neq 0$ and $b \neq 0$.

P.56. For the first example, let $\Omega = [0,1]$ with the Borel σ-algebra and measure. Then take $X_n(\omega) = \mathbf{1}_{[0,1/n]}(\omega)$.

Let us make an example where the limit does not holds. We take all the X_n independent (we know we can do it with a product Ω). Notice that $\lim_n X_n(\omega) = 0$ iff $X_n(\omega) = 0$ for all $n \geq N(\omega)$. This means that $\{\lim_n X_n = 0\} = \cup_{m \geq 1} A_m$ where $A_m = \cap_{j \geq m} \{X_j = 0\}$. However by independence

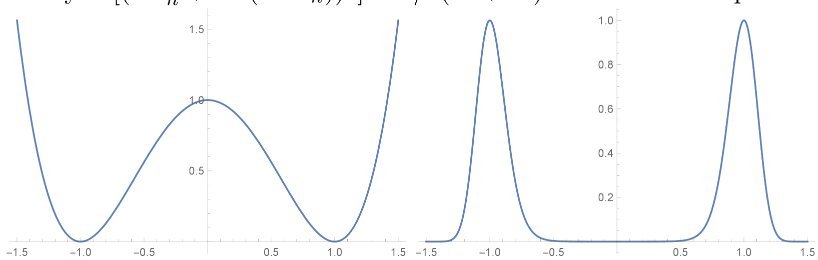
$$\mathbb{P}(A_m) = \prod_{j \geq m} \mathbb{P}(X_j = 0) = \prod_{j \geq m} (1 - \tfrac{1}{j}) = 0$$

So $\mathbb{P}(A_m) = 0$ and $\mathbb{P}(\cup_m A_m) \leq \sum_m \mathbb{P}(A_m) = 0$. Therefore $\mathbb{P}(\lim_n X_n = 0) = 0$.

P.57. $\lim_n X_n(x)$ does not exists for a.e. $x \in [0,1]$, as indeed $X_n(x)$ takes infinitely many times but the value 0 and 1 (even if 1 happens more and more rarely as n increases). On the other hand $\mathbb{P}(X_n > \varepsilon) = 2^{-h_n} \to 0$. Finally, $2^{h_n} \leq n$ so the series diverges. We could have deduced the divergence from the Borelli-Cantelli lemma and the first two remark.

P.58. $X_n = nY_n$ where Y_n is Bernoulli of parameter ε_n. Then $\mathbb{M}[|X_n|^q] = n^q \varepsilon_n$. For the first question we choose $\varepsilon_n = cn^{-p}$, for the second $\varepsilon_n = n^{-p} \log n$. For the third question, $\varepsilon_n = n^{-p}/\log n$.

P.59. For any continuous function f with at most polynomial growth, the expectation $\mathbb{M}[f(X_n)]$ in a (Riemann) integral $\int \varrho_n(x) f(x) dx$. When n is large, it is easily seen that, for each $\delta > 0$, the integration away from $[-1-\delta, -1+\delta] \cup [1-\delta, 1+\delta]$ vanishes exponentially in n. So we have $\lim_n \mathbb{M}[f(X_n)] = 1/2(f(-1) + f(+1))$. We deduce that $\lim_n \mathbb{M}[3X_n^2 + \cos(2\pi X_n)] = (3 + \cos(-2\pi) + 3 + \cos(2\pi))/2 = 4$. Similarly $\mathbb{M}[(3X_n^2 + \cos(2\pi X_n))^2] = 1/2(4^2 + 4^2) = 16$. So the required limit is 0.



P.60. We have that $\mathbb{P}(X_n = k)$ converges to the respective probability for the Poisson distribution. Indeed from Euler limit $(1 + z_n/n) \to e^z$ and Stirling formula (k fixed, n to ∞)

$$\begin{aligned}\binom{n}{k} p_n^k (1-p_n)^{n-k} &= \frac{n!}{n^k k!(n-k)!} (np_n)^k (1 - (np_n)/n)^n (1 - (np_n)/n)^{-k} \\ &= \frac{e^{-\lambda}\lambda^k}{k!} \frac{n!}{n^k (n-k)!} (1 + o(1)) = \frac{e^{-\lambda}\lambda^k}{k!}\end{aligned}$$

From this, the limit of $\mathbb{M}[f(X_n)] = \sum_k \binom{n}{k} p_n^k (1 - p_n^k) f(k)$ converges to the Poisson limit for each compactly supported function f. So we gather convergence in distribution.

P.61. Here we compute the characteristic function. Let $U_n := \frac{1}{n} \sum_{j=1}^n a_j X_j$. Then

$$\begin{aligned}\varphi_n(t) :=& \mathbb{M}[e^{itU_n}] = \mathbb{M}[\prod_{j=1}^n e^{ita_j X_j/n}] = \prod_{j=1}^n \mathbb{M}[e^{ita_j X_j/n}] = \prod_{j=1}^n (1-p_j) * 1 + p_j e^{ita_j/n} \\ & \prod_{j=1}^n (1 + p_j(e^{ita_j/n} - 1)) = \prod_{j=1}^n (1 + itp_j a_j/n)(1 + o(1))\end{aligned} \tag{13}$$

Now, recall the Euler limit: if $z_n \to z$ in $\mathbb{C}$, then $(1 + z_n/n)^n \to e^z$. Therefore, passing to the limit in (13), $\lim_n \varphi(t) = e^{itc}$ if and only if $p_j a_j \to c$.

So the sequence converges iff $p_j a_j$ as a limit. As the sequence is bounded, it will converge at least along subsequences.

P.62. Notice that $Z_{2n} = \sqrt{1/2}Z_n + \sqrt{1/2}W_n$, where Z_n and W_n are independent and $\mathcal{N}(0,1)$. If Z_n converges to some Z in probability, then so does Z_{2n}, and Z is necessarily $\mathcal{N}(0,1)$. Now however, $W_n = \sqrt{2}Z_{2n} - Z_n \to (\sqrt{2}-1)Z$ which is not possible as again $W_n \sim \mathcal{N}(0,1)$.

P.63. Actually we have seen this during the lectures. (We have seen that, since the limit is continuous, $\mathbb{P}(X_n \in I) \to \mathbb{P}(X \in I)$ for any interval I, and quickly mentioned uniformity). It is enough to prove the uniform convergence of the distribution function $F_n(x) := \mathbb{P}(X_n \le x)$.

Fix $\varepsilon > 0$. We want to find N_ε such that $|F_n(x) - F(x)| \le \varepsilon$ for all $n \ge N_\varepsilon$ and all $x \in \mathbb{R}$.

- Fix $L \equiv L_\varepsilon$ such that $\mathbb{P}(|X| \ge L) \le \varepsilon/3$.
- Since $(-L, L)^c$ is closed, $\lim_n \mathbb{P}(|X_n| \ge L) \le \mathbb{P}(|X| \ge L)$, and therefore for some N'_ε and $n \ge N_\varepsilon$, it also holds $\mathbb{P}(|X_n| \ge L) \le \varepsilon/2$. Therefore for $x < -L$, $|F_n(x) - F(x)| \le \varepsilon/2 + \varepsilon/3 \le \varepsilon$. Similarly for $x > L$, $|F_n(x) - F(x)| = |\mathbb{P}(X_n > x) - \mathbb{P}(X > x)| \le \varepsilon$.
- We know that $F_n \to F$ pointwise. However, since F is continuous, this convergence is uniform on compact sets. So for some N''_ε and $n \ge N_\varepsilon$, $|F_n(x) - F(x)| \le \varepsilon/2$ for all $x \in [-L, L]$.
- For $N_\varepsilon = \max(N'_\varepsilon, N''_\varepsilon)$ both work.

P.64. Let X_i be the result of the i-th dice. Then $\mathbb{M}[X_i] =: m = 7/2$ and $\mathbb{D}(X_i) =: s = 35/12$. The X_i are i.i.d. random variables and setting $Z_n = (\sum_i (X_i - m))/\sqrt{sn}$ we have for

$$\mathbb{P}(\sum_i X_i > l) = \mathbb{P}(Z_n > \frac{l - nm}{\sqrt{sn}})$$

Now Z_n goes in distribution to $\mathcal{N}(0,1)$ and $a := \frac{l-nm}{\sqrt{sn}} = 2$. Therefore

$$\mathbb{P}(\sum_i X_i \ge l) = \mathbb{P}(Z \ge a)(1 + O(n^{-1/2}))$$

and $\mathbb{P}(Z \ge a) = \int_a^\infty e^{-x^2/2}(2\pi)^{-1/2}dx \simeq 0.022750132$. The error is actually bounded (uniformly in a) by $1.4 * 10^{-9}$, so all the digits in this estimate are significant.

P.65. It is the same computation as before, but now the unknown is n. If X_i corresponds to the success in the i-th try, we have that, for

$$Z_n := \sum_i (X_i - p)/\sqrt{np(1-p)} = \sum_{i=1}^n \frac{X_i - \mathbb{M}[X_i]}{\sqrt{n\mathbb{D}(X_i)}}$$

$$\mathbb{P}(\sum_i X_i > l) = \mathbb{P}(Z_n > (l - np)/\sqrt{np(1-p)}) \simeq \int_{(l-np)/\sqrt{np(1-p)})}^\infty e^{-x^2/2}(2\pi)^{-1/2}dx$$

Let $-r$ be the value for which the Gaussian integral from $-r$ to ∞ is 0.95. (Any statistician knows by hearth that $r \simeq 1.645$). So we need to solve for n

$$(l - np)/\sqrt{np(1-p)}) = -r$$

For $l = 100$ and $p = 10^{-5}$ we get $n \sim 3273.8$, so we need to rent it for 3274 hours. Notice that the exact calculation would give the same number, but we would need to rent the same computer to find it out.

P.66. Take $p_{n,k} = \delta_{k,n}$. The limit is not a probability. However, if we know that $\mathbb{M}[\phi(X_n)] \le c$ for some $\phi \ge 0$ such that $\phi(x) \to +\infty$ when $x \to \infty$, then we can reason as in the Bernoulli to Poisson exercises, to find that X_n converges.