Абсолютно неперервні випадкові величини: щільність та математичне сподівання.

1 Теоретичні відомості

Що ж це таке, абсолютно неперервна випадкова величина?

Нехай $(\Omega, \mathcal{F}, P)$ – імовірнісний простір. Введемо на просторі випадкову величину ξ та через $F_\xi(t)$ позначимо її функцію розподілу.

Якщо існує така вимірна функція $f_\xi(t)$, що F_ξ допускає представлення

$$F_\xi(t) = \int\limits_{-\infty}^{t} f_\xi(u)du,\ t \in \mathbb{R},$$

тоді ξ називають абсолютно неперервною випадковою величиною.

Більш загально, f_ξ є похідною Радона-Никодима міри Лебега-Стілтьєса, породженою $F_X(\cdot)$, відносно міри Лебега на прямій λ_1:

$$F_X(A) = \int\limits_A f_\xi(u)\lambda_1(du),\ A \in \mathcal{B}(\mathbb{R}).$$

Взагалі кажучи, можна розширити поняття щільності для довільної сигма-скінченної міри λ, а не зосереджуватися конкретно на λ_1 (тоді і для дискретних розподілів знайдеться відповідна міра та 'щільність').

Щільність розподілу $f_\xi(t)$ має дві основні властивості:

1. Невід'ємність: $f_\xi(t) \geq 0$ для всіх $t \in \mathbb{R}$,
2. Нормованість: $\int\limits_{-\infty}^{+\infty} f_\xi(u)du = 1$.

Неважко переконатися в тому, що для абсолютно неперервної в.в. ξ: $F_\xi(t)$ неперервна по t.

Математичне сподівання для абсолютно неперервної випадкової величини можна обчислити наступним чином:

$$E[\xi] = \int\limits_{-\infty}^{+\infty} t f_\xi(t)dt,$$

а LOTUS перезапишеться аналогічно (див. попереднє заняття):

$$E[g(\xi)] = \int\limits_{-\infty}^{+\infty} g(t) f_\xi(t)dt.$$

Приклади абсолютно неперервних розподілів

Назва розподілу	**Параметри**	**Позначення**	**Щільність,** $f_\xi(t)$
Рівномірний	$a, b \in \mathbb{R}$ $a < b$	$\xi \sim U[a,b]$	$1_{[a,b]}(t)\dfrac{1}{b-a}$
Розподіл трикутника	$a, b, c \in \mathbb{R}$ $a \le c \le b,\ a < b$	$\xi \sim \mathrm{Triag}(a,b,c)$	$\left(1_{(a,c]}(t)\cdot\dfrac{t-a}{c-a} + 1_{(c,b)}(t)\cdot\dfrac{b-t}{b-c}\right)\cdot\dfrac{2}{b-a}$
Експоненційний	$\lambda > 0$	$\xi \sim \mathrm{Exp}(\lambda)$	$1_{(0,\infty)}(t)\lambda e^{-\lambda t}$
Гамма-розподіл	$\alpha > 0,\ \lambda > 0$	$\xi \sim \Gamma(\alpha,\lambda)$	$1_{(0,+\infty)}(t)\dfrac{1}{\Gamma(\alpha)}t^{\alpha-1}\lambda^{\alpha-1}\lambda e^{-\lambda t}$
Бета-розподіл	$\alpha > 0,\ \beta > 0$	$\xi \sim \mathrm{B}(\alpha,\beta)$	$1_{(0,1)}(t)\dfrac{1}{\mathrm{B}(\alpha,\beta)}t^{\alpha-1}(1-t)^{\beta-1}$
Нормальний розподіл	$\mu \in \mathbb{R},\ \sigma > 0$	$\xi \sim N(\mu,\sigma^2)$	$\dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\dfrac{(t-\mu)^2}{2\sigma^2}\right)$
Логнормальний розподіл	$\mu \in \mathbb{R},\ \sigma > 0$	$\xi \sim LN(\mu,\sigma^2)$	$1_{(0,+\infty)}(t)\dfrac{1}{\sqrt{2\pi\sigma^2}t}\exp\left(-\dfrac{(\ln(t)-\mu)^2}{2\sigma^2}\right)$
Розподіл хі-квадрат	$k \in \mathbb{N}$	$\xi \sim \chi^2_k$	$1_{(0,+\infty)}(t)\dfrac{1}{\Gamma(k/2)}t^{k/2-1}(1/2)^{k/2-1}(1/2)e^{-(1/2)t}$
Розподіл Коші	-	$\xi \sim Cauchy$	$\dfrac{1}{\pi}\dfrac{1}{1+t^2}$

Декілька коментарів:

1. Взагалі рівномірний розподіл можна узагальнити, про це неявно йшла мова в занятті з геометричної імовірності.

2. Деякі розподіли, можуть подавати в іншій параметризації, виходячи з певних інтерпретацій. Наприклад, у цій постановці подано гамма-розподіл з параметром інтенсивності $\lambda > 0$. В іншій літературі замість параметра інтенсивності може бути введено параметр масштабу: $\theta := 1/\lambda$.

3. Так, деякі розподіли, справді кажучи, є частковими випадками більш загальних розподілів (хі-квадрат, експоненційний якраз і є гамма-розподілами, або ж стандартний рівномірний розподіл є бета розподілом).

4. Розподіл випадкової величини легше задати щільністю (якщо така існує), коли функція розподілу має складну форму (не можна явно виразити).

   Також буде зручно подавати розподіли в термінах перетворень від випадкових величин. Цими перетвореннями будемо займатися найближчим часом.

З допитливості можете спробувати самостійно познаходити для наведених вище розподілів $E[\xi]$, $Var[\xi]$. Можна також спробувати відшукати *медіану* розподілу: таку точку t_*, щоб $P(\xi < t_*) = 1/2$. Також можна графіки нарисувати.

2 Задачі

2.1 Задача 1

Нехай $\xi \sim \text{Exp}(\lambda)$, $\lambda > 0$.

1. Знайти функцію розподілу ξ.
2. Обчислити 'хвіст' розподілу $P(\xi \geq t)$, обчислити $P(\xi \geq t + s \mid \xi \geq s)$ для $t, s > 0$.
3. Обчислити $E[\xi^k]$, $k > 0$. Звідси отримати $E[\xi]$ та $Var[\xi]$.
4. Знайти розподіл випадкових величин $\eta = \exp(a\xi)$, $|a| > 0$, та $\nu = [\xi]$, де $[x]$ – ціла частина числа x.

Розв'язання

Оскільки ξ має експоненційний розподіл, то за означенням її щільність $f_\xi(t) = 1_{(0,\infty)}(t)\lambda e^{-\lambda t}$. Отже,

$$F_\xi(t) = \int\limits_{-\infty}^{t} f_\xi(u)du = \int\limits_{-\infty}^{t} 1_{(0,\infty)}(u)\lambda e^{-\lambda u}du = \begin{cases} 0, & t \leq 0, \\ -e^{-\lambda u}\Big|_0^t = 1 - e^{-\lambda t}, & t > 0. \end{cases}$$

Знайдемо хвіст розподілу ξ, перейшовши до доповнення події $\{\xi \geq t\}$:

$$P(\xi \geq t) = 1 - P(\xi < t) = 1 - F_\xi(t) = e^{-\lambda t},\ t > 0.$$

Тепер можна легко обчислити умовну імовірнсть:

$$P(\xi \geq t + s \mid \xi \geq s) = \frac{P(\xi \geq t + s, \xi \geq s)}{P(\xi \geq s)} = \frac{P(\xi \geq t + s)}{P(\xi \geq s)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t},\ t, s > 0.$$

Тобто експоненційний розподіл має властивість відсутності пам'яті.

Для $k > 0$ знайдемо $E[\xi^k]$, скориставшись LOTUS:

$$E[\xi^k] = \int\limits_{-\infty}^{+\infty} t^k f_\xi(t)dt = \int\limits_{0}^{+\infty} t^k \lambda e^{-\lambda t}dt = \lambda^{-k}\int\limits_{0}^{+\infty} (\lambda t)^k e^{-\lambda t}d(\lambda t) = \lambda^{-k}\int\limits_{0}^{+\infty} u^{(k+1)-1}e^{-u}du = \lambda^{-k}\Gamma(k+1).$$

Зокрема $E[\xi] = \lambda^{-1}$ та $E[\xi^2] = 2\lambda^{-2}$, звідки $Var[\xi] = E[\xi^2] - (E[\xi])^2 = \lambda^{-2}$ (можете звісно перевірити, взявши інтеграли в $E[\xi]$ та $E[\xi^2]$ частинами).

Тепер знайдемо розподіл $\eta = e^{a\xi}$, $|a| > 0$. Для $a > 0$

$$P(\eta < t) = P(e^{a\xi} < t) = P(a\xi < \ln(t)) = P(\xi < \ln(t)/a) = \begin{cases} 0, & t \leq 1, \\ 1 - e^{-(\lambda/a)\ln(t)} = 1 - t^{-(\lambda/a)}, & t > 1. \end{cases}$$

Навпаки, для $a < 0$

$$P(\eta < t) = P(a\xi < \ln(t)) = P(\xi > \ln(t)/a) = \begin{cases} 0, & t \le 0, \\ e^{-(\lambda/a)\ln(t)} = t^{-(\lambda/a)}, & 0 < t \le 1, \\ 1, & t > 1. \end{cases}$$

Зокрема з цього можна побачити, що при $a = -\lambda$ маємо, що $\eta \sim U[0,1]$ (по суті частковий випадок квантильного перетворення).

Тепер перевіримо, що ж то за розподіл для $\nu = [\xi]$. Очевидно, що ν є дискретною випадковою величиною з носією на $\mathbb{Z}_+$, тому досить знайти імовірності набувать значень в цих точках

$$\begin{aligned} P(\nu = k) = P([\xi] = k) = P(k \le \xi < k+1) = P(\xi < k+1) - P(\xi < k) = \\ = F_\xi(k+1) - F_\xi(k) = e^{-\lambda(k+1)} - e^{-\lambda k} = q^k(1-q),\ k \in \mathbb{Z}_+, \end{aligned}$$

де $q = e^{-\lambda}$. Тобто $\nu \sim \mathrm{Geom}(1-q)$.

2.2 Задача 2

Нехай $\xi \sim U[0,1]$, $-\infty < a < b < +\infty$.

1. Обчислити $E[1/(\xi+1)]$, $E[\sin(2\pi\xi)]$.
2. Знайти розподіл випадкової величини $\eta = (b-a)\xi + a$, $a \neq 0$, $a < b$.
3. Знайти розподіл випадкової величини $\nu = 3|\xi - 2|$.
4. Обчислити $E[\nu]$, $Var[\nu]$.

Розв'язання

Оскільки ξ має рівномірний розподіл, то за означенням її щільність $f_\xi(t) = 1_{(0,1)}(t)$.

Підрахунок математичних сподівань очевидний з LOTUS:

$$E[1/(\xi+1)] = \int\limits_0^1 \frac{1}{x+1}dx = \ln(x+1)\Big|_0^1 = \ln(2).$$

$$E[\sin(2\pi\xi)] = \int\limits_0^1 \sin(2\pi x)dx = -\cos(2\pi x)/(2\pi)\Big|_0^1 = 0.$$

Знайдемо розподіл лінійного перетворення $\eta = (b-a)\xi + a$:

$$P(\eta < t) = P((b-a)\xi < t - a) = F_\xi((t-a)/(b-a)),\ t \in \mathbb{R}$$

Звідси $f_\eta(t) = (P(\eta < t))'_t = f_\xi((t-a)/(b-a))/(b-a) = 1_{(a,b)}(t)/(b-a)$, тобто $\eta \sim U[a,b]$.

Знайдемо тепер розподіл $\nu = 3|\xi - 2|$. Для $t > 0$

$$P(\nu < t) = P(2 - t/3 < \xi < 2 + t/3) = P(\xi < 2 + t/3) - P(\xi < 2 - t/3) = 1 - P(\xi < 2 - t/3).$$

Знайдемо щільність ν:

$$f_\nu(t) = \frac{1}{3}f_\xi(2 - t/3) = \frac{1}{3}1_{(0,1)}(2 - t/3) = \frac{1}{3}1_{(3,6)}(t).$$

Випадок $t \leq 0$ очевидний. Отже $\nu \sim U[3,6]$ та (див. попереднє заняття)

$$E[\nu] = (3+6)/2 = 9/2,\ Var[\nu] = (6-3)^2/12 = 9/12 = 3/4.$$

2.3 Задача 3

Абсолютно неперервна випадкова величина ξ має симетричний розподіл, якщо розподіли ξ та $-\xi$ співпадають.

Сформулювати умови симетричності розподілу:

1. В термінах функції розподілу ξ,
2. В термінах щільності розподілу ξ.

Розв'язання

Позначимо через $F_\xi(t) = P(\xi < t)$ – функцію розподілу ξ. Якщо $\xi =^d -\xi$, то це значить, що

$$P(\xi < t) = P(-\xi < t),\ t \in \mathbb{R}.$$

З іншого боку, $P(-\xi < t) = P(\xi > -t) = 1 - P(\xi \leq -t) = 1 - P(\xi < -t)$, тобто

$$F_\xi(t) = 1 - F_\xi(-t),$$

отримавши умову на функцію розподілу. Далі, беремо похідну по t:

$$f_\xi(t) = 0 + f_\xi(-t) = f_\xi(-t),$$

тобто f_ξ має бути парною. Отримали умову на щільність розподілу.

2.4 Задача 4

Нехай випадкова величина $\xi \sim Cauchy$.

1. Чи є ξ інтегровною?
2. Знайти розподіл $\eta = 1/\xi$.

Розв'язання

Розглянемо $\xi = \min(0, \xi) + \max(0, \xi) =: \xi_- + \xi_+$.

Якщо доведемо інтегровність $\xi_\pm$, тоді інтегровною буде ξ (чому?).

Спочатку дослідимо ξ_+: згідно LOTUS

$$E[\xi_+] = E[0 \cdot 1_{0 \geq \xi} + \xi \cdot 1_{\xi > 0}] = E[\xi 1_{\xi > 0}] = \frac{1}{\pi} \int\limits_0^{+\infty} \frac{t}{1+t^2} dt$$

Інтеграл у правій частині рівності розбіжний за ознакою еквівалентності:

$$\frac{t}{1+t^2} \sim \frac{1}{t},\ t \to +\infty,$$

де $\int_0^{+\infty} (1/t) dt = +\infty$, отже $E[\xi_+] = +\infty$. Аналогічно можна показати, що $E[\xi_-] = -\infty$.

Виходить, що $E[\xi]$ не існує, отже випадкова величина ξ не є інтегровною.

Знайдемо розподіл $1/\xi$. Неважко переконатися, що при $t < 0$

$$\begin{aligned} P(1/\xi < t) &= P(1/\xi < t, \xi < 0) + P(1/\xi < t, \xi > 0) = P(1/\xi < t, \xi < 0) = \\ &= P(\xi > 1/t, \xi < 0) = P(1/t < \xi < 0) = \\ &= F_\xi(0) - F_\xi(1/t) \\ &\Rightarrow f_{1/\xi}(t) = (F_\xi(0) - F_\xi(1/t))'_t = 0 - \frac{1}{\pi} \frac{1}{1+1/t^2} (1/t^2) = \frac{1}{\pi} \frac{1}{1+t^2}. \end{aligned}$$

Далі, при $t > 0$:

$$\begin{aligned} P(1/\xi < t) &= 1 - P(1/\xi \geq t) = 1 - (P(1/\xi \geq t, \xi < 0) + P(1/\xi \geq t, \xi > 0)) = \\ &= 1 - P(1/\xi \geq t, \xi > 0) = 1 - P(\xi \leq 1/t, \xi > 0) = 1 - F_\xi(1/t) + F_\xi(0) \\ &\Rightarrow f_{1/\xi}(t) = (1 - F_\xi(1/t) + F_\xi(0))'_t = 0 - \frac{1}{\pi} \frac{1}{1+1/t^2} (1/t^2) + 0 = \frac{1}{\pi} \frac{1}{1+t^2}. \end{aligned}$$

Отже $f_{1/\xi}(t) = f_\xi(t)$ м.с. λ_1, тобто $1/\xi$ теж має розподіл Коші.

В принципі можна отримати потрібний результат в термінах $F_{1/\xi}(t)$, скориставшись властивістю $\arctan(t) + \arctan(1/t) = \mathrm{sign}(t) \cdot \pi/2$.

2.5 Задача 5

Випадкова величина ξ має щільність $f_\xi(x)$, а функція $g : \mathbb{R} \to \mathbb{R}$ кусково строго монотонна і кусково неперервно диференційовна. Доведіть, що щільність розподілу випадкової величини $\eta = g(\xi)$ має вигляд

$$f_\eta(y) = \sum_{x:g(x)=y} \frac{f_\xi(x)}{|g'(x)|}$$

Розв'язання

Через I_k, $k \in I$ (I – зліченна множина) позначимо інтервали монотонності g, $\cup_k I_k = \mathbb{R}$. Тоді

$$P(\eta < t) = P(g(\xi) < t) = \sum_{k \in I} P(g(\xi) < t, \xi \in I_k)$$

Через g_k позначимо звуження функції g на $I_k := (a_k, b_k]$ (відповідно при $b_k = +\infty$ правий кінець без включення).

Нехай на проміжку I_k функція g строго зростає по t і $\{x \in I_k \mid g(x) < t\} = (a_k, g_k^{-1}(t)) \subset I_k$.

$$P(g(\xi) < t, \xi \in I_k) = P(\xi < g_k^{-1}(t), \xi \in I_k) = \int\limits_{a_k}^{g_k^{-1}(t)} f_\xi(u)du = \left| z := g(u), u = g_k^{-1}(z), du = \frac{dz}{g_k'(g_k^{-1}(z))} \right| =$$

$$= \int\limits_{g(a_k)}^{t} \frac{f_\xi(g_k^{-1}(z))}{g_k'(g_k^{-1}(z))} dz \Rightarrow \frac{d}{dt} P(g(\xi) < t, \xi \in I_k) = \frac{f_\xi(g_k^{-1}(t))}{g_k'(g_k^{-1}(t))} = \frac{f_\xi(g_k^{-1}(t))}{|g_k'(g_k^{-1}(t))|}.$$

Навпаки, припустимо що на I_k функція g строго спадає по t та $\{x \in I_k \mid g(x) < t\} = (g_k^{-1}(t), b_k] \subset I_k$. Тоді

$$P(g(\xi) < t, \xi \in I_k) = P(\xi > g_k^{-1}(t), \xi \in I_k) = P(g_k^{-1}(t) < \xi \le b_k) = \int\limits_{t}^{g_k(b_k)} \frac{f_\xi(g_k^{-1}(z))}{g_k'(g_k^{-1}(z))} dz$$

$$\Rightarrow \frac{d}{dt} P(g(\xi) < t, \xi \in I_k) = -\frac{f_\xi(g_k^{-1}(t))}{g_k'(g_k^{-1}(t))} = \frac{f_\xi(g_k^{-1}(t))}{|g_k'(g_k^{-1}(t))|}.$$

А тепер припустимо, що для всіх $x \in I_k$ виконується нерівність $g(x) < t$. Тоді

$$P(g(\xi) < t, \xi \in I_k) = P(\xi \in I_k) \Rightarrow \frac{d}{dt} P(g(\xi) < t, \xi \in I_k) = 0.$$

Залишається врахувати випадок, коли на I_k нерівність $g(x) < t$ не справджується для всіх $x \in I_k$. Тоді події в імовірності несумісні, звідси

$$P(g(\xi) < t, \xi \in I_k) = 0 \Rightarrow \frac{d}{dt} P(g(\xi) < t, \xi \in I_k) = 0.$$

В результаті

$$f_\eta(t) = \sum_{k \in I} 1_{g_k((a_k, b_k])}(t) \frac{f_\xi(g_k^{-1}(t))}{|g_k'(g_k^{-1}(t))|} = \sum_{y:g(y)=t} \frac{f_\xi(y)}{|g'(y)|},$$

де $g_k(\pm\infty) := \lim\limits_{x \to \pm\infty} g_k(x)$.