Całka Lebesgue'a musi odejść?

Rafał SZTENCEL *

– Ależ proszę pana, nie ma czegoś takiego, jak przeciętny czas operacji.

Serial „Na dobre i na złe", odc. 197 (12 XII 2004)

Wygląda na to, że scenarzyści (lub konsultanci) serialu, wkładając w usta siostry Marty to stwierdzenie, zdawali sobie sprawę, że wartość oczekiwana może nie istnieć.

Wracamy do wzoru z poprzedniego odcinka:

$$(*) \qquad EX = \int_0^\infty P(X > t)\,\mathrm{d}t,$$

gdzie X jest nieujemną zmienną losową. Pozbędziemy się tego ograniczenia.

Niech X będzie dowolną zmienną losową o dystrybuancie F. Definiujemy jej część dodatnią i ujemną:

$$X^+ = \max(X, 0), \quad X^- = \max(-X, 0) = -\min(X, 0).$$

Jasne jest, że tak określone zmienne losowe są nieujemne, ponadto

$$X = X^+ - X^-, \quad |X| = X^+ + X^-.$$

Jeśli $t \geqslant 0$, to

$$P(X^+ > t) = P(X > t), \quad P(X^- > t) = P(X < -t).$$

Na mocy wzoru $(*)$ mamy:

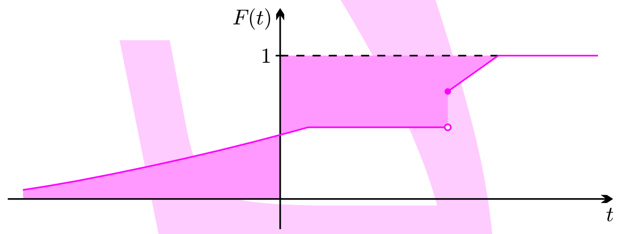
$$\begin{aligned} EX &= E(X^+ - X^-) = EX^+ - EX^- = \\ &= \int_0^\infty P(X > t)dt - \int_0^\infty P(X < -t)dt = \\ &= \int_0^\infty P(X > t)dt + \int_{-\infty}^0 P(X < t)dt = \\ &= \int_0^\infty P(X > t)dt + \int_{-\infty}^0 P(X \leqslant t)dt = \\ &= \int_0^\infty (1 - F(t))dt + \int_{-\infty}^0 F(t)dt. \end{aligned}$$

Być może komentarza wymaga równość

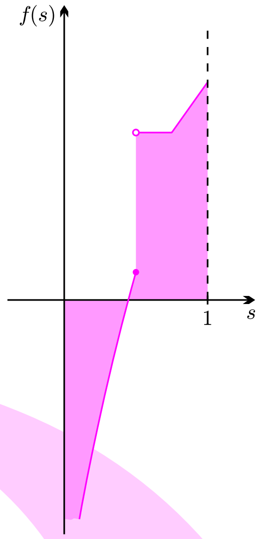
$$\int_{-\infty}^0 P(X < t)dt = \int_{-\infty}^0 P(X \leqslant t)dt.$$

Czytelnik zechce się zastanowić, dlaczego różnica powyższych całek, czyli $\int_{-\infty}^0 P(X = t)dt$ jest zerem. Swoją drogą, to bardzo wygodne, bo na przykład wzór $(*)$ można pamiętać tylko z grubsza.

Udowodnione właśnie uogólnienie tego wzoru zilustrujemy na rysunku.



A teraz zamienimy rolami osie układu współrzędnych. Wykres dystrybuanty F zostanie wówczas przekształcony przez symetrię względem prostej o równaniu $y = x$. Otrzymana krzywa nie zawsze jest wykresem funkcji, czemu łatwo w razie potrzeby zaradzić – rysunek pokazuje, że wystarczy usunąć pionowe kreski (powstałe z przedziałów stałości dystrybuanty), a luki (powstałe ze skoków dystrybuanty) wypełnić poziomymi kreskami. Okazuje się wtedy, że $EX = \int_0^1 f(s)\mathrm{d}s$, gdzie f jest (z grubsza) funkcją odwrotną do dystrybuanty F.



Istnieje ścisły przepis na f:

$$f(s) = \sup\{u: F(u) < s\}, \quad s \in (0, 1).$$

Z własności dystrybuanty łatwo wynika, że f jest dobrze określona, bowiem zbiór $A_s = \{u: F(u) < s\}$ jest przedziałem otwartym postaci $(-\infty, t)$, gdzie oczywiście $f(s) = t$. Może wygodniej stwierdzić, że zbiór $A'_s = \{u: F(u) \geqslant s\}$ jest domknięty: wynika to z prawostronnej ciągłości F. Niech $t = \inf A'_s$. Jeśli $F(u_n) \geqslant s$, $u_n \geqslant t$, $n = 1, 2, \ldots$, $u_n \to t$, to $F(t) \geqslant s$. Dlatego $t \in A'_s$.

Funkcja f jest określona na odcinku $(0, 1)$. Jeśli za prawdopodobieństwo P uznamy miarę Lebesgue'a λ na tym odcinku, to f stanie się zmienną losową. Jaka jest dystrybuanta f? Zobaczmy.

Mamy ciąg równoważności:

$$t \geqslant f(s) \equiv t \geqslant \sup A_s \equiv t \notin A_s \equiv s \leqslant F(t).$$

Wobec tego

$$P(f(s) \leqslant t) = \lambda((0, F(t))) = F(t).$$

Z przeprowadzonego powyżej rozumowania wynikają zapewne rozmaite wnioski – praktyczne i teoretyczne. Praktyczny jest taki: dość łatwo napisać program, generujący liczby losowe o rozkładzie jednostajnym na przedziale $(0, 1)$. Jeśli chcemy mieć liczby losowe o rozkładzie wyznaczonym przez dystrybuantę F, wystarczy skonstruować uogólnioną funkcję odwrotną do f (może Czytelnicy zorientowali się, że $f(s)$ jest jej s-tym kwantylem).

A teoretyczny? Wartość oczekiwaną dowolnej zmiennej losowej umiemy obliczyć, całkując dość regularną, bo niemalejącą i lewostronnie ciągłą funkcję f. Do tego celu nie potrzeba wcale miary i całki Lebesgue'a. A prawdopodobieństwo na $(0, 1)$? Ostatecznie można powiedzieć parę słów o prawdopodobieństwie geometrycznym i wystarczy. Czy studenci, których celem jest uzyskanie licencjatu, naprawdę muszą zapoznawać się z tak trudną teorią? O tym za miesiąc.

*Instytut Matematyki Uniwersytetu Warszawskiego