解答例

STEP III 2015、第 13 問

確率・統計 ・ 一様分布に従う確率変数の関数

STEP III 2015、問 13

先に進む前に、まず問題を読んでおこう。問題は STEP Support Programme の問題データベースにある：

step.maths.org/questions/15-s3-q13

鍵となる考え方

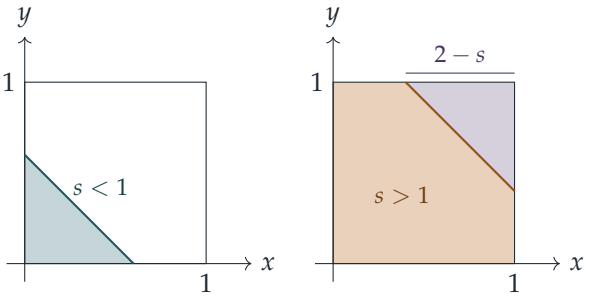
X と Y は独立（independent）で、$[0,1]$ 上の一様分布（uniform distribution）に従うので、組 (X,Y) は単位正方形（unit square）上に一様に分布する。したがって、X と Y に関する確率はすべて単に**面積**である。$X+Y \leq s$ や $Y \leq rX$ のような事象は正方形を直線で切り分け、その片側の面積は三角形の面積、または正方形から三角形を除いた部分の面積になる。そのあと、$1/(X+Y)$ と $X/(X+Y)$ の分布は、それらの事象を、すでに面積を求めた事象を用いて書き直すことで得られる。

解答

以下では一貫して、(X,Y) は面積 1 の単位正方形 $[0,1]^2$ 上の一様分布に従う。よって正方形内の任意の領域 A について、$\mathrm{P}((X,Y) \in A)$ は A の面積に等しい。

小問 (i)

$S = X+Y$ とおくと、これは $[0,2]$ の値をとる。また $F(s) = \mathrm{P}(S \leq s)$ とおく。事象 $S \leq s$ は、正方形のうち直線 $x+y=s$ 上またはその下側にある部分である。



$0 \le s \le 1$ のとき、この直線は座標軸と $(s,0)$ および $(0,s)$ で交わり、その下側の領域は直角をはさむ 2 辺の長さが s の直角二等辺三角形である。$1 \le s \le 2$ のときは、直線の**上側**の領域が、角 $(1,1)$ にある三角形になる。直角をはさむ 2 辺は直線から角までのびており、その長さは $1-(s-1)=2-s$ である。よって

$$F(s) = \mathrm{P}(X+Y \le s) = \begin{cases} 0 & s < 0, \\ \frac{1}{2}s^2 & 0 \le s \le 1, \\ 1 - \frac{1}{2}(2-s)^2 & 1 \le s \le 2, \\ 1 & s > 2. \end{cases}$$

中央の 2 つの式は $s=1$ で一致する（どちらも $\frac{1}{2}$ を与える）。当然そうでなければならない。

小問 (ii)

$T = 1/S$ とおく。$0 < S \le 2$ であるから（値 $S=0$ をとる確率はゼロである）、T は $[\frac{1}{2}, \infty)$ の値をとる。$t > 0$ に対して、$s \mapsto 1/s$ は $s > 0$ で減少するから、

$$\mathrm{P}(T \le t) = \mathrm{P}\left(S \ge \tfrac{1}{t}\right) = 1 - F\left(\tfrac{1}{t}\right),$$

ここで、S は連続型であり、したがって $\mathrm{P}(S = 1/t) = 0$ となることを用いた。

> **考えてみよう** F のどちらの式を用いるのか。それは t の位置ではなく、$1/t$ の位置によって決まる。

$t \ge 1$ ならば $0 < 1/t \le 1$ なので、1 つ目の式を用いる：$\mathrm{P}(T \le t) = 1 - \frac{1}{2}t^{-2}$。$\frac{1}{2} \le t \le 1$ ならば $1 \le 1/t \le 2$ なので、2 つ目の式を用いる：$\mathrm{P}(T \le t) = \frac{1}{2}\left(2 - t^{-1}\right)^2$。また、$t < \frac{1}{2}$ に対しては $\mathrm{P}(T \le t) = 0$ である。

微分すると確率密度関数（density）が得られる。$t \ge 1$ のとき、

$$f(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(1 - \tfrac{1}{2}t^{-2}\right) = t^{-3},$$

また $\frac{1}{2} \leq t \leq 1$ のときは、連鎖律（chain rule）により

$$f(t) = \left(2 - t^{-1}\right) \cdot t^{-2} = 2t^{-2} - t^{-3},$$

であり、それ以外では $f(t) = 0$ である。これが求める確率密度関数である。

期待値（expectation）は次のようになる。

$$\begin{aligned} \mathrm{E}\left(\frac{1}{X+Y}\right) = \int_{1/2}^{\infty} t f(t)\,\mathrm{d}t &= \int_{1/2}^{1} \left(2t^{-1} - t^{-2}\right)\mathrm{d}t + \int_{1}^{\infty} t^{-2}\,\mathrm{d}t \\ &= \left[2\ln t + t^{-1}\right]_{1/2}^{1} + \left[-t^{-1}\right]_{1}^{\infty} \\ &= (0+1) - (-2\ln 2 + 2) + 1 \\ &= 2\ln 2. \end{aligned}$$

t^{-2} は $[1,\infty)$ で可積分であるから、この広義積分（improper integral）は収束する。

小問 (iii)

$R = Y/X$ とおくと、これは $[0,\infty)$ の値をとる（事象 $X = 0$ の確率はゼロである）。$r \geq 0$ に対して、事象 $R \leq r$ は $Y \leq rX$ であり、これは正方形のうち、原点を通る直線 $y = rx$ 上またはその下側にある部分である。

$0 \leq r \leq 1$ のとき、この直線は右側の辺上の点 $(1, r)$ で正方形から出ていき、その下側の領域は底辺 1、高さ r の三角形で、面積は $\frac{1}{2}r$ である。$r \geq 1$ のときは、直線は上側の辺上の点 $(1/r, 1)$ で正方形から出ていき、その**上側**の領域は、上側の辺に沿った底辺 $1/r$、高さ 1 の三角形で、面積は $\frac{1}{2r}$ である。よって

$$H(r) = \mathrm{P}\left(\frac{Y}{X} \leq r\right) = \begin{cases} 0 & r < 0, \\ \frac{1}{2}r & 0 \leq r \leq 1, \\ 1 - \dfrac{1}{2r} & r \geq 1. \end{cases}$$

次に $Z = \dfrac{X}{X+Y} = \dfrac{1}{1+R}$ とおくと、これは $[0,1]$ の値をとる。$0 < z \leq 1$ に対して、$r \mapsto 1/(1+r)$ は $r \geq 0$ で減少するから、

$$\mathrm{P}(Z \leq z) = \mathrm{P}\left(1 + R \geq \frac{1}{z}\right) = \mathrm{P}\left(R \geq \frac{1-z}{z}\right) = 1 - H\left(\frac{1-z}{z}\right).$$

ここでも、どちらの式を用いるかは $(1-z)/z$ の位置によって決まる。それが 1 以下となるのは、ちょうど $z \geq \frac{1}{2}$ のときである。

- $\frac{1}{2} \leq z \leq 1$ のとき：$\mathrm{P}(Z \leq z) = 1 - \dfrac{1-z}{2z} = \dfrac{3}{2} - \dfrac{1}{2z}$。

- $0 < z \le \frac{1}{2}$ のとき：$\mathrm{P}(Z \le z) = 1 - \left(1 - \dfrac{z}{2(1-z)}\right) = \dfrac{z}{2(1-z)}$。

どちらも $z = \frac{1}{2}$ で $\frac{1}{2}$ となる。微分すると（2 つ目については $\frac{z}{2(1-z)} = \frac{1}{2}\left(\frac{1}{1-z} - 1\right)$ と書き直す）、Z の確率密度関数は

$$g(z) = \begin{cases} \dfrac{1}{2(1-z)^2} & 0 \le z \le \frac{1}{2}, \\ \dfrac{1}{2z^2} & \frac{1}{2} \le z \le 1, \\ 0 & \text{それ以外.} \end{cases}$$

$g(1-z) = g(z)$ であることに注意しよう。すなわち、この確率密度関数は $\frac{1}{2}$ に関して対称である。

期待値。 これは積分せずに書き下すことができる。X と Y は独立で同じ分布に従うので、組 (X, Y) と (Y, X) は同じ分布をもつ。よって $\dfrac{X}{X+Y}$ と $\dfrac{Y}{X+Y}$ は同じ期待値をもつ。それらの和はちょうど 1 であるから、

$$\mathrm{E}\left(\frac{X}{X+Y}\right) = \tfrac{1}{2}.$$

積分による確認。

$$\mathrm{E}(Z) = \int_0^{1/2} \frac{z}{2(1-z)^2}\,\mathrm{d}z + \int_{1/2}^{1} \frac{1}{2z}\,\mathrm{d}z.$$

2 つ目の積分は $\frac{1}{2}\left[\ln z\right]_{1/2}^{1} = \frac{1}{2}\ln 2$ である。1 つ目の積分では $u = 1 - z$ と置換（substitution）する。すると $\mathrm{d}z = -\mathrm{d}u$ であり、$z = 0, \frac{1}{2}$ は $u = 1, \frac{1}{2}$ に対応する。

$$\begin{aligned} \int_0^{1/2} \frac{z}{2(1-z)^2}\,\mathrm{d}z &= \frac{1}{2}\int_{1/2}^{1} \frac{1-u}{u^2}\,\mathrm{d}u = \frac{1}{2}\left[-\frac{1}{u} - \ln u\right]_{1/2}^{1} \\ &= \frac{1}{2}\left(-1 - (-2 + \ln 2)\right) = \frac{1}{2}(1 - \ln 2). \end{aligned}$$

足し合わせると $\mathrm{E}(Z) = \frac{1}{2}(1 - \ln 2) + \frac{1}{2}\ln 2 = \frac{1}{2}$ となり、上で求めた値が確かめられる。

検算

- どの分布関数（distribution function）も 0 から 1 まで増加し、式のつなぎ目で連続である：$F(1) = \frac{1}{2}$、$\mathrm{P}(T \le 1) = \frac{1}{2}$、$H(1) = \frac{1}{2}$、$\mathrm{P}(Z \le \frac{1}{2}) = \frac{1}{2}$。これらがいずれも $\frac{1}{2}$ となるのは、正方形が対角線に関して対称だからである。
- 確率密度関数 f を積分すると 1 になる：$\int_{1/2}^{1}(2t^{-2} - t^{-3})\,\mathrm{d}t = 2 - \frac{3}{2} = \frac{1}{2}$ かつ $\int_1^{\infty} t^{-3}\,\mathrm{d}t = \frac{1}{2}$。

- $2\ln 2 \approx 1.386$ は当然 1 と 2 の間にあり、また凸関数（convex function）$1/s$ に対するイェンセン（Jensen）の不等式から予想されるとおり、$1/\mathrm{E}(X+Y) = 1$ より大きい。正方形上での $\iint (x+y)^{-1}\,\mathrm{d}x\,\mathrm{d}y$ を直接数値計算すると 1.3863 となる。分布関数および $\mathrm{E}(Z) = \frac{1}{2}$ は、コンピュータシミュレーションによっても確認した。

よくある間違い

- F や H の式を、$1/t$ や $(1-z)/z$ ではなく t や z の値に基づいて選んでしまい、2 つの式を取り違える。
- $1/s$ や $1/(1+r)$ が減少関数であり、不等号の向きを逆にしなければならないことを忘れる。$T \leq t$ は $S \geq 1/t$ を意味する。
- 小問 (i) の角にある三角形の、直角をはさむ辺の長さを間違える。それは $s-1$ ではなく $2-s$ である。

STEP のその先へ

(X, Y) に関する命題を面積に置き換え、それから分布関数を微分するという方法は、確率変数の関数の分布を求めるための一般的な手法である**分布関数法**（distribution function method）である。大学では、これに同時確率密度関数に対する変数変換の公式（change-of-variables formula、ヤコビアンを用いる）が加わり、$X+Y$ と $X/(X+Y)$ を同時に扱うことができる。特に、ここでの $X+Y$ の確率密度関数は、2 つの一様分布の確率密度関数の**畳み込み**（convolution）である。Grimmett と Welsh の *Probability: An Introduction* は両方の方法を扱っており、Blitzstein と Hwang の *Introduction to Probability* には同じ趣旨の例が数多く載っている。

Beyond STEP 解答集：https://beyondstep.keithmcnulty.org/ja/。解答本文は CC0 1.0 である。問題そのものは © Cambridge University Press & Assessment に帰属し、ここには掲載していない。英語版から機械翻訳したものであり、まだ校閲を経ていない。