A Distribution of Conditional Convolution of Random Variables: "Individual" Relative Deprivation

Hiroshi Hamada
Tohoku University

2014/10/16

1　問題

y が他者の所得，x が自分 i の所得であると仮定する．このとき個人 i の剥奪度は

$$z_i = \begin{cases} y - x, & x < y \\ 0, & x \geq y \end{cases}$$

の平均値である (Yitzhaki 1979; Hey and Lambert 1980)．Yitzhaki が平均値として個人剥奪度を定義したことからも明らかなように，ある特定の個人の感じる剥奪感 z_i にもばらつきがある (それは 0 や 3.25 や 10.078 といった様々な値をとりうる．ただし負の値は定義上存在しない)．

いま，抽象的に他者の所得が確率変数 Y, 自己の所得が確率変数 Y で表されていると仮定する．このとき合成

$$z = \begin{cases} y - x, & x < y \\ 0, & x \geq y \end{cases}$$

によって作られた Z の確率分布は，個人が感じる剥奪感の強度を，社会全体でまとめた分布である．Z は数学的には $Z = Y - X(y > x$ のとき) とおいて合成した条件付きの確率変数である．

Yitzhaki の個人剥奪度が，自己の所得を固定したうえで，他者所得だけがランダムに変動した場合の自己と他者の差の (条件付きの) 期待値であるのに対して，確率変数 Z は，期待値を取る前の個人毎に異なる剥奪感の分布を，社会全体で一つに合成した分布である．

Yitzhaki の個人剥奪度が，平均という操作によって個人の剥奪体験（これ自体が一つの分布を持つ）を特定の実現値として要約した値であるのに対して，確率変数 Z は，様々な個人が感じた剥奪度を，同じ強度であるもの同士をまとめて，社会全体で一つの分布として表したものである．

「条件付きの合成積」という操作の応用的価値を探るために，以下，平均化する前の剥奪感を社会レベルで総合した確率変数 Z の明示的表現について考える.

確率変数 Z の分布の具体的イメージを次のような数値計算例 (Mathematica) で示す.

```
Module[{x, d},
 x = RandomVariate[NormalDistribution[3, 1], 30];
```

```
 d = Flatten[Table[ If[x[[i]] < x[[j]], x[[j]] - x[[i]], 0],
{j, 1, Length[x]}, {i, 1, Length[x]}]];
 Histogram[d] ]
```

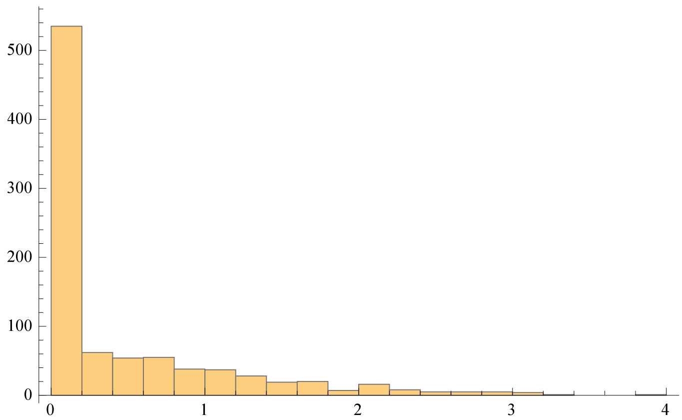

図 1: 確率変数 Z の分布,$N(3,1), n = 30$ の数値例

2 モデル

もし Z が条件分岐しない関数であれば,

$$\begin{cases} z = y - x \\ w = y \end{cases}$$

とおいて，確率変数のたたみ込みから確率密度関数を特定すればよい．すなわち

$$\begin{cases} x = w - z \\ y = w \end{cases}$$

として，$f(x, y)$ を X, Y の同時確率密度関数，Z の確率密度関数を $p(z)$ とおけば

$$p(z) = \int_{-\infty}^{\infty} f(w - z, w) \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} dw$$

である (小針 1973:100-103).

さて今考えているケースは，確率変数 Z が条件分岐するという特殊なケースであり．この特殊条件を考慮しながら Z の分布（確率密度関数 $p(z)$）を特定しなければならない．考えるべき変数変換は

$$\begin{cases} z = \begin{cases} y - x, & x < y \\ 0, & x \geq y \end{cases} \\ w = y \end{cases}$$

すなわち

$$\begin{cases} x = \begin{cases} w - z, & z > 0 \\ w, & z \leq 0 \end{cases} \\ y = w \end{cases}$$

である．

準備として X, Y の分布は独立同分布だから，それぞれの確率密度関数を $f(x), f(y)$ とおけば，常に

$$\int\int_{y \geq x} f(x)f(y)dxdy = \frac{1}{2}$$

が成立することを確認しておく．

命題 1 (独立同分布の差の分布の一性質)**.** 確率変数 X, Y は独立であり, 同じ分布（確率密度関数）$f(x), f(y)$ を持つと仮定する．このとき

$$\int\int_{y \geq x} f(x)f(y)dxdy = \frac{1}{2}$$

証明.

$$\begin{aligned}
\int\int_{y \geq x} f(x)f(y)dxdy &= \int_{\infty}^{\infty} \left(\int_{y}^{\infty} f(x)f(y)dx \right) dy \\
&= \int_{-\infty}^{\infty} f(y) \left(\int_{y}^{\infty} f(x)dx \right) dy \\
&= \int_{-\infty}^{\infty} f(y)\{1 - F(y)\}dy \\
&= \int_{-\infty}^{\infty} f(y)dy - \int_{-\infty}^{\infty} f(y)F(y)dy \\
&= 1 - \int_{-\infty}^{\infty} f(y)F(y)dy \\
&= 1 - \frac{1}{2} \qquad (\text{部分積分を使った})
\end{aligned}$$

□

このとき次が成立する．

命題 2 (個人剥奪度 Z の確率密度関数)**.** 確率変数 X が自己の所得, Y が他者の所得を表し，独立に同分布に従うと仮定する．このとき (平均をとる前の) 個人剥奪度 Z

$$z = \begin{cases} y - x, & x < y \\ 0, & x \geq y \end{cases}$$

の分布 $p(z)$ は

$$p(z)=\begin{cases}0, & z<0\\ 1/2, & z=0\\ \displaystyle\int_{-\infty}^{\infty} f(w-z,w)\begin{vmatrix}\frac{\partial x}{\partial z} & \frac{\partial x}{\partial w}\\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w}\end{vmatrix} dw, & z>0\end{cases}$$

に従う.

証明. まず Z が条件分岐しない場合にその確率密度関数 $p(z)$ の意味を確認しておく. X, Y を合成して作った Z の確率密度関数は，特定の条件を満たす $f(x,y)$ の断面積 (周辺分布) であると考えられる (小針 1973:102-103).

例えば，$Z=Y-X$ という関係があるとき，同時確率密度関数 $f(x,y)$ の $z=y-x$ という線上の断面を考える．するとこの断面の面積 S は $p(z)$ に一致する．というのも断面の面積 S は

$$S=\int_{-\infty}^{\infty} f(x,z+x)dx=\int_{-\infty}^{\infty} f(w-z,w)dw=p(z)$$

だからである ($x=w-z$ とおいて変数変換した．ヤコビアンは $dx/dw=1$).

$y-x=z>0$(言い換えれば $y>x$) の場合，求める分布 $p(z)$ は, 変数変換定理より, $x, y, f(x,y)$ 軸の 3 次元空間における, $y-x=z>0$ を満たす $f(x,y)$ の断面積 $p(z)$ と一致する.

一方，$y-x=z\leq 0$(言い換えれば $y\leq x$) であるような $f(x,y)$ の断面積 $p(z)$ の場合には，相対的剥奪の定義より $z<0$ ならば常に $z=0$ となることに注意する.

$y-x=\leq 0$ であるとき，断面 $p(z)$ の $z\in(-\infty,0]$ の範囲における総和は

$$\int_{-\infty}^{0} p(z)dz$$

である．ところで先に示した命題 1 より

$$\int_{0}^{\infty} p(z)dz=\int\int_{y\geq x} f(x,y)dxdy=\frac{1}{2}$$

であり

$$\int_{-\infty}^{\infty} p(z)dz=1$$

という確率の仮定から

$$\int_{-\infty}^{0} p(z)dz=\frac{1}{2}$$

でなくてはならない.

一方，$z<0$ の範囲では z は常に 0 であり，負の z は実現値が空だから，

$$P(Z<0)=P(\emptyset)=0 \quad (\text{測度 } 0)$$

と考えるのが自然である.

つまり確率密度関数 $p(z)$ は $z<0$ の範囲ではずっと 0 で $z=0$ のとき 1/2 にジャンプし，$z>0$ の範囲で $p(z)$ になるような関数である．このように定義すれば

$$\int_{-\infty}^{\infty} p(z)dz = 1$$

である．すなわち明示的には

$$p(z) = \begin{cases} 0, & z<0 \\ 1/2, & z=0 \\ \displaystyle\int_{-\infty}^{\infty} f(w-z,w) \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} dw, & z>0 \end{cases}$$

である． □

命題 2 より，$z<0$ の範囲で確率密度が 0，$z=0$ の点で確率密度が 1/2 という事実は，確率変数 X,Y の分布がどのようなものであれ成立する．

簡単な例を示そう．

例 1. $X \sim N(\mu,\sigma^2), Y \sim N(\mu,\sigma^2)$ のとき，相対的剥奪度 $Z=Y-X$(ただし $x>y$ ならば $z=0$) の分布は

$$p(z) = \begin{cases} 0, & z<0 \\ 1/2, & z=0 \\ \dfrac{1}{\sqrt{2\pi 2\sigma^2}} \exp\left\{-\dfrac{z^2}{4\sigma^2}\right\}, & z>0 \end{cases}$$

である．

証明．正規分布の再生性より $z=y-x$ の分布は $Z \sim N(0,2\sigma^2)$ にしたがう．その確率密度関数は

$$\frac{1}{\sqrt{2\pi 2\sigma^2}} \exp\left\{-\frac{z^2}{4\sigma^2}\right\}$$

である．あとは $x>y$ ならば $z=0$ という条件を考慮して命題 2 を適用すればよい． □

X,Y が対数正規分布に従う場合，残念ながら「差をとる」という合成に関して再生性が成立しないので $Z=Y-X$ の確率密度関数を明示的に示すことはできない（積分が綺麗に解けない）が，近似計算によっておおよその個人剥奪度の分布を理論的に推定することができる．

また命題 2 より，次が成立する．

命題 3. 剥奪度 Z のレンジを等分したとき，0 を含む区間が最頻値となる．つまり任意の資源分布に関して，剥奪度 Z が 0 付近の実現確率が最も多い．

証明．剥奪度 Z のレンジを $[0,\varepsilon),[\varepsilon,2\varepsilon),[2\varepsilon,3\varepsilon),[3\varepsilon,4\varepsilon)...$ のように等間隔に分割する．その範囲で剥奪度 Z が実現する確率を計算すると，

$$P(0 \le Z < \varepsilon) = P(0) + P(0<Z<\varepsilon) = \frac{1}{2} + \alpha, \qquad (\alpha > 0)$$

一方，ε 以上となる全ての区間を積分しても，1/2 に満たない．よって区間 $[0, \varepsilon)$ で Z が実現する確率が最大である． □

この命題の妥当性を，数値計算によって確認する．以下は対数正規分布に従う確率変数を条件付きで合成した確率変数の分布である．

```
Module[{x, d},
 x = RandomVariate[LogNormalDistribution[3, 1], 30];
 d = Flatten[Table[ If[x[[i]] < x[[j]], x[[j]] - x[[i]], 0],
{j, 1, Length[x]}, {i, 1, Length[x]}]];
 Histogram[d] ]
```

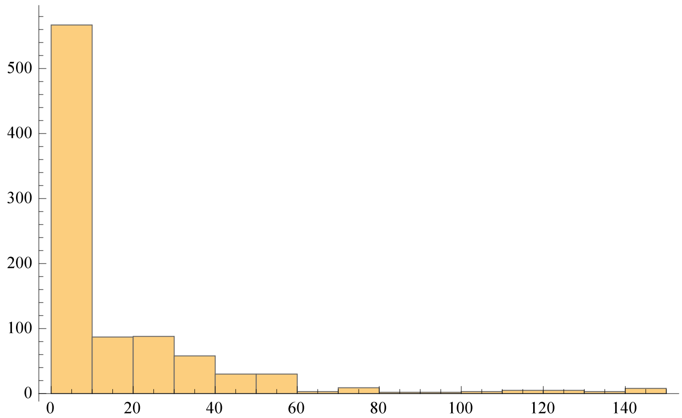

図 2: 確率変数 Z の分布,$\Lambda(3,1), n = 30$ の数値例

数値計算例から，資源分布の形状に関わりなく，0 付近が個人剥奪度 Z のモードとなることが確認できた．

[文献]

D'Ambrosio, Conchita, and Joachim R. Frick, 2007, "Income satisfaction and relative deprivation: an empirical link," *Social Indicators Research*, 81(3): 497-519.

浜田宏, 2001,「経済的地位の自己評価と準拠集団 —— δ 区間モデルによる定式化」『社会学評論』206 号（第 52 巻・第 2 号）: 283-99.

Hey, John. D. and Peter J. Lambert, 1980, "Relative deprivation and the Gini coefficient: comment," *Quarterly Journal of Economics*, 95(3): 567-573.

石田淳, 2011,「相対的剥奪と準拠集団の計量モデル」『理論と方法』26(2): 371-88.

Yitzhaki, Shlomo, 1979, "Relative deprivation and the Gini coefficient," *Quarterly Journal of Economics*, 93(2): 321-324.