解答例

STEP III 2012、第 13 問

確率・統計 ・ 切断正規分布

STEP III 2012、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/12-s3-q13

鍵となる考え方

区間上の条件付き期待値（conditional expectation）は、その区間上での $z\,f(z)$ の積分を、その区間の確率で割ったものである。標準正規分布（standard Normal distribution）では、$z\,\mathrm{e}^{-z^2/2}$ はちょうど $-\mathrm{e}^{-z^2/2}$ の導関数になっているので、この積分は厳密に計算できる。あとは $X = \mu + \sigma Z$ と標準化（standardising）し、X の符号に応じて $|X|$ を分けて考えるだけである。

解答

以下、Z の確率密度関数（probability density function）を $\phi(z) = \dfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-z^2/2}$ と書く。すると $\Phi'(z) = \phi(z)$ かつ $\phi'(z) = -z\,\phi(z)$ である。

小問 (i)

$a < Z < b$ という条件のもとでの Z の条件付き密度は、(a, b) 上で $\phi(z)/\mathrm{P}(a < Z < b)$ であり、$\mathrm{P}(a < Z < b) = \Phi(b) - \Phi(a)$ である。よって

$$\begin{aligned}
\mathrm{E}(Z \mid a < Z < b) &= \frac{1}{\Phi(b) - \Phi(a)} \int_a^b z\,\phi(z)\,\mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}\,(\Phi(b) - \Phi(a))} \Big[-\mathrm{e}^{-z^2/2} \Big]_a^b \\
&= \frac{\mathrm{e}^{-a^2/2} - \mathrm{e}^{-b^2/2}}{\sqrt{2\pi}\,(\Phi(b) - \Phi(a))},
\end{aligned}$$

となる。ここで $\dfrac{\mathrm{d}}{\mathrm{d}z}(-\mathrm{e}^{-z^2/2}) = z\,\mathrm{e}^{-z^2/2}$ を用いた。これが求める結果である。積分は収束するので、$\mathrm{e}^{-\infty} = 0$、$\Phi(\infty) = 1$ と読めば、$b = \infty$（あるいは $a = -\infty$）の場合にもこの式は成り立つ。

小問 (ii)

$Z \sim \mathrm{N}(0,1)$、$\sigma > 0$ として $X = \mu + \sigma Z$ と書く。$\sigma > 0$ であるから

$$X > 0 \iff Z > -\mu/\sigma,$$

であり、$X > 0$ という条件をつけることは $Z > -\mu/\sigma$ という条件をつけることと同じである。期待値の線形性（linearity）は条件付きでも成り立つので

$$\mathrm{E}(X \mid X > 0) = \mathrm{E}(\mu + \sigma Z \mid Z > -\mu/\sigma) = \mu + \sigma\,\mathrm{E}(Z \mid Z > -\mu/\sigma),$$

となり、示すべき式が得られる。

$|X|$ の平均。 $c = \mu/\sigma$ とおく。小問 (i) で $a = -c$、$b = \infty$ とすると

$$\mathrm{E}(Z \mid Z > -c) = \frac{\mathrm{e}^{-c^2/2}}{\sqrt{2\pi}\,(1 - \Phi(-c))} = \frac{\phi(c)}{p}, \qquad \text{ただし} p = \mathrm{P}(X > 0) = 1 - \Phi(-c).$$

直前の式の両辺に p を掛けると

$$\mathrm{E}(X;\, X > 0) := \mathrm{E}(X \mid X > 0)\,\mathrm{P}(X > 0) = \mu p + \sigma\phi(c).$$

> **考えてみよう** 2 つ目の積分を計算せずに、$X < 0$ の部分からの寄与を求めるにはどうすればよいか？

2 つの部分を足し合わせると全体の平均になる：$\mathrm{E}(X;\, X > 0) + \mathrm{E}(X;\, X < 0) = \mathrm{E}(X) = \mu$（事象 $X = 0$ の確率は 0 である）。したがって $\mathrm{E}(X;\, X < 0) = \mu - \mu p - \sigma\phi(c)$ である。$X < 0$ のとき $|X| = -X$ であるから

$$\begin{aligned} m = \mathrm{E}|X| &= \mathrm{E}(X;\, X > 0) - \mathrm{E}(X;\, X < 0) \\ &= 2(\mu p + \sigma\phi(c)) - \mu \\ &= \mu(2p - 1) + 2\sigma\phi(c). \end{aligned}$$

最後に $2p - 1 = 1 - 2\Phi(-\mu/\sigma)$、$2\phi(c) = \sqrt{2/\pi}\,\mathrm{e}^{-\mu^2/(2\sigma^2)}$ を用いると

$$m = \mu(1 - 2\Phi(-\mu/\sigma)) + \sigma\sqrt{2/\pi}\,\exp(-\tfrac{1}{2}\mu^2/\sigma^2),$$

となり、示すべき式が得られる。

$|X|$ の分散。要点は $|X|^2 = X^2$ であることで、二次モーメント（second moment）を求めるのに新たな計算は要らない：

$$\mathrm{E}(|X|^2) = \mathrm{E}(X^2) = \mathrm{Var}(X) + (\mathrm{E}X)^2 = \sigma^2 + \mu^2.$$

したがって

$$\mathrm{Var}|X| = \mu^2 + \sigma^2 - m^2.$$

検算

- $\mu = 0$ のとき、$|X|$ は半正規分布（half-Normal distribution）に従い、公式から $m = \sigma\sqrt{2/\pi}$ というよく知られた値が得られる。このとき分散は $\sigma^2(1 - 2/\pi)$ であり、正である。
- $\mu \gg \sigma$ のとき、X はほぼ常に正であり、$\Phi(-\mu/\sigma) \approx 0$ で指数関数の項もごく小さいので、$m \approx \mu$、$\mathrm{Var}|X| \approx \sigma^2$ となる。これは当然そうなるべき結果である。
- m の公式は μ を $-\mu$ に置き換えても変わらない（$1 - 2\Phi(-c) = -(1 - 2\Phi(c))$ を用いる）。これは対称性 $|X| \sim |-X|$ と合致している。
- 数値的には、$\mu = 1$、$\sigma = 2$ のとき $m \approx 1.7912$、$\mathrm{Var}|X| \approx 1.7917$ となり、直接の数値積分（quadrature）や 2×10^5 個の値によるシミュレーションの結果と一致する。

よくある間違い

- 小問 (i) で $\Phi(b) - \Phi(a)$ で割るのを忘れること。積分だけでは $\mathrm{E}(Z;\, a < Z < b)$ であり、条件付き期待値ではない。
- $\mathrm{E}|X| = \mathrm{E}(X \mid X > 0)$ と書いてしまうこと。条件付き期待値には $\mathrm{P}(X > 0)$ と $\mathrm{P}(X < 0)$ で重みをつけなければならない。
- $|X|^2 = X^2$ からただちに得られるのに、$\mathrm{E}(X^2)$ を別の積分で計算してしまうこと。

STEP のその先へ

$|X|$ の分布は**折り畳み正規分布**（folded Normal distribution）と呼ばれ、小問 (i) の条件付き分布は**切断正規分布**（truncated Normal distribution）である。比 $\phi(c)/\big(1 - \Phi(-c)\big)$ は逆ミルズ比（inverse Mills ratio）と呼ばれ、データが打ち切られたり（censored）選択されたりする計量経済学（econometrics）のモデルで中心的な役割を果たす。また、$\phi'(z) = -z\phi(z)$ という工夫は、正規近似（Normal approximation）のためのスタイン

（Stein）の方法の出発点でもある。条件付き期待値についてさらに学ぶには、Blitzstein and Hwang, *Introduction to Probability*（確率論入門）を参照するとよい。

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。