現代数理統計学の基礎
～第 2 章 問 6～

担当：???

2021 年 4 月 22 日

概要

生存関数の足し上げは期待値という問題。生存時間解析の授業で連続変数の場合の証明はしたものの、離散型変数の場合はしたことがなかったため、難しく感じた。教科書通りの解法と、わかりやすい解法と、連続変数の場合の解法を載せる。

1 問題文

問 6. X を非負の整数上で確率をもつ離散型確率変数とし，非負の整数 k に対して $F(k) = P(X \leq k)$ とする。X の期待値が存在するとき次の等式が成り立つことを示せ。

$$E[X] = \sum_{k=0}^{\infty}(1 - F(k))$$

2 準備

2.1 期待値

定義 2.8（p.17）より、問題文のような確率変数 X の期待値が存在するとき、$E[X] < \infty$ である。また、同定義より、X が離散型変数で確率関数が $f(x)$ の時には、普通は、

$$E[X] = \sum_{k=0}^{\infty} kf(k) \quad (\because \text{非負の離散値をとるので } k = 0 \text{ 始まり})$$

である。よって、この問題は

$$\sum_{k=0}^{\infty} kf(k) = \sum_{k=0}^{\infty}(1 - F(k))$$

を、示せば良い。

2.2 累積分布関数 cdf について

定義 2.2（p.12）より、非負の整数上で確率をもつ離散型確率変数 X の累積分布関数が $F_X(x)$ で表され（しばしば下付きの X は省略される）、確率関数が $f(x)$ で表される時、

$$
\begin{aligned}
F_X(x) &= P(X \le x) \\
&= \sum_{i=0}^{x} f(i) \quad (\because \text{非負の離散値をとるので } i = 0 \text{ 始まり})
\end{aligned}
$$

である。ここで、全ての実現値に対応する全ての確率の和は 1 になるように定義されているので、$F_X(\infty) = \sum_{i=0}^{\infty} f(i) = 1$ となる。これより、(1) の右辺のシグマの中は、

$$
\begin{aligned}
1 - F(k) &= F(\infty) - F(k) \\
&= \sum_{i=0}^{\infty} f(i) - \sum_{i=0}^{k} f(i) \\
&= \sum_{i=k+1}^{\infty} f(i)
\end{aligned}
$$

となる。これを回答の最後で使う。

2.3 二重シグマについて

一般的な二重シグマは交換が容易にできるが、左側のシグマの足し上げる変数が右側のシグマの上端に含まれている場合は交換は工夫が必要である。求める無限級数が絶対収束することが分かっているとき、

$$
\sum_{i=0}^{\infty} \sum_{j=0}^{i} f(i, j) = \sum_{j=0}^{\infty} \sum_{i=j}^{\infty} f(i, j)
$$

が成り立つ。私はこの問題の素晴らしい証明方法を思いついたが、それを書くにはこの余白は狭すぎるので下の**普通の回答**で感覚を感じて欲しい。

3 回答

3.1 普通の回答

X の確率関数を $f(x)$, 分布関数を $F(x)$ と表すと、題意より、

$$
\begin{aligned}
E[X] &= \sum_{k=0}^{\infty} k f(k) \\
&= 1 \times f(1) + 2 \times f(2) + 3 \times f(3) + 4 \times f(4) + \cdots \\
&= \begin{array}{ccccccccc}
f(1) & + & f(2) & + & f(3) & + & f(4) & + & \cdots \\
 & + & f(2) & + & f(3) & + & f(4) & + & \cdots \\
 & & & + & f(3) & + & f(4) & + & \cdots \\
 & & & & & + & f(4) & + & \cdots
\end{array} \\
&= P(X \ge 1) + P(X \ge 2) + P(X \ge 3) + P(X \ge 4) + \cdots \\
&= \sum_{k=0}^{\infty} P(X \ge k) \\
&= \sum_{k=0}^{\infty} (1 - F(k))
\end{aligned}
$$

3.2 本に載っているようなきちんとした回答

まず、X は非負の実現値をとる離散型確率変数であるから、その期待値が存在するとき、X の確率関数を $f(x)$ と表すと、

$$E[X] = \sum_{k=0}^{\infty} kf(k)$$

ここで、$k = \sum_{n=0}^{k-1} 1$ であるから、

$$\begin{aligned}
E[X] &= \sum_{k=0}^{\infty} kf(k) \\
&= \sum_{k=0}^{\infty} (\sum_{n=0}^{k-1} f(k)) \\
&= \sum_{n=0}^{\infty} \sum_{k=n+1}^{\infty} f(k) \\
&= \sum_{n=0}^{\infty} (\sum_{k=0}^{\infty} f(k) - \sum_{k=0}^{n} f(k)) \\
&= \sum_{n=0}^{\infty} (1 - \sum_{k=0}^{n} f(k)) \\
&= \sum_{n=0}^{\infty} (1 - F(n)) \\
&= \sum_{k=0}^{\infty} (1 - F(k))
\end{aligned}$$

3.3 非負の連続型変数の時の回答

$$\int_0^{\infty} 1 - F(x)dx = \int_0^{\infty} xf(x)dx$$

を示す。まず左辺は、

$$\begin{aligned}
\int_0^{\infty} 1 - F(x)dx &= \int_0^{\infty} (x)' \times \big\{1 - F(x)\big\}dx \\
&= \big[x\big\{1 - F(x)\big\}\big]_0^{\infty} - \int_0^{\infty} x \times \frac{d}{dx}\big\{1 - F(x)\big\}dx
\end{aligned}$$

ここで、$\frac{d}{dx}\big\{1 - F(x)\big\} = 0 - \frac{d}{dx}F(x) = -f(x)$ であるから、

$$\int_0^{\infty} 1 - F(x)dx = \big[x\big\{1 - F(x)\big\}\big]_0^{\infty} + \int_0^{\infty} xf(x)dx \tag{1}$$

よって題意より、

$$\begin{aligned}
&\big[x\big\{1 - F(x)\big\}\big]_0^{\infty} = 0 \\
\Rightarrow &\lim_{x \to \infty} x\big\{1 - F(x)\big\} = 0
\end{aligned}$$

を示せば良い。

ここで、

$$
\begin{aligned}
x\big\{1-F(x)\big\} &= x\int_x^{\infty} f(k)dk \\
&\leq k\int_x^{\infty} f(k)dk \\
&= \int_x^{\infty} kf(k)dk
\end{aligned}
$$

ゆえに、これを代入したい。いま、$E[X]=\int_0^{\infty} xf(x)dx<\infty$ であるから、

$$
\begin{aligned}
\lim_{x\to\infty} x\big\{1-F(x)\big\} &\leq \lim_{x\to\infty}\int_x^{\infty} kf(k)dk \\
&= 0
\end{aligned}
$$

したがって、式 (1) より、

$$
\int_0^{\infty} 1-F(x)dx=\int_0^{\infty} xf(x)dx=E[X]
$$

が成り立つ。