10月1日「データ解析のための数理統計入門」第4章
4-3 条件付き期待値前半

2024-10-01

目次

条件付き確率分布に関する期待値を条件付き期待値と言います。

定義 4.9（条件付き期待値）

二つの確率変数X, Yと関数$g(X)$を考える。

(A)離散確率変数の場合

$Y = y_i$ を与えた時のXの条件付き確率関数 $p_{X \mid Y}(x_j \mid y_i)$ で表すと、$Y = y_i$を与えた時の$g(X)$の条件付き期待値は

$$E[g(X) \mid Y = y_i] = \sum_{j=1}^{\infty} g(x_j) p_{X \mid Y}(x_j \mid y_i)$$

(B)連続確率変数の場合

$Y = y$ を与えた時のXの条件付き確率密度関数 $f_{X \mid Y}(x \mid y)$ で表すと、$Y = y$ を与えた時の $g(X)$ の条件付き期待値は

$$E[g(X) \mid Y = y] = \int_{-\infty}^{\infty} g(x_j) f_{X \mid Y}(x \mid y) dx$$

$Y = y$ を与えた時のXの条件付き平均$E[X \mid Y = y]$ 、条件付き分散 $Var[X \mid Y = y] = E[(X - E[X|Y = y])^2|Y = y]$ と表される。

周辺分布の平均と分散を求める際、条件付き平均や条件付き分散を求める方が易しくなる場合がある。 →5 章ででてくるらしいです。

命題 4.10 条件付き期待値の性質

⑴繰り返し期待値の法則

$$E[X] = E[E[X \mid Y]]$$

右辺については、まず$E[X \mid Y]$を求め、次にyをYで置き換えてYに関して期待値を取ることを意味する。

⑴証明

連続分布の場合、右辺

おそらく離散分布の場合は右辺は下のようになる。