情報通信システム(特)論II

Information and Communication Systems II

第1回講義資料

Lecture notes 1

確率論の復習

A Review of Probability Theory

豊橋技術科学大学

Toyohashi University of Technology

電気・電子情報工学系

Department of Electrical and Electronic Information Engineering

准教授　竹内啓悟

Associate Professor　Keigo Takeuchi

確率変数(Random variables)

離散確率変数(Discrete random variable)

確率変数Xは離散的な値$x_i \in \mathcal{X}$(例えば、$\mathcal{X} = \{0, 1\}$)を取る。

A random variable X takes discrete values $x_i \in \mathcal{X}$ (e.g. $\mathcal{X} = \{0, 1\}$).

$P(X = x_i) = p_i$: 確率$p_i \in [0, 1]$で確率変数Xは値x_iを取る。

The random variable X takes x_i with probability $p_i \in [0, 1]$.

連続確率変数(Continuous random variable)

$p_X(x)$を連続確率変数Xの確率密度関数とする。

Let $p_X(x)$ denote the probability density function (pdf) of a continuous random variable X.

確率変数Xが区間$(a, b]$に入る確率

Probability with which the random variable X is in the interval $(a, b]$.

$$P(X \in (a, b]) = \int_a^b p_X(x) dx\,.$$

期待値(Expectation)

任意の関数fに対して、(For any function f)

離散確率変数X (Discrete random variable)

$$\mathbb{E}[f(X)] = \sum_i f(x_i)P(X = x_i).$$

連続確率変数X (Continuous random variable)

$$\mathbb{E}[f(X)] = \int_{-\infty}^{\infty} f(x)p_X(x)dx.$$

期待値は、実現値の総和を取って、その総数で割ったものではない。

The expectation is not the sum of realizations divided by the number of realizations.

n次元離散確率ベクトル(n-dimensional discrete random vector)

同時確率分布(Joint probability distribution) $P(\boldsymbol{X}=\boldsymbol{x})=p_{\boldsymbol{x}}$

離散確率ベクトル$\boldsymbol{X}=(X_1,\dots,X_n)^{\mathrm{T}}$はベクトル$\boldsymbol{x}\in\mathcal{X}^n$を確率$p_{\boldsymbol{x}}$で取る。

A discrete random vector $\boldsymbol{X}=(X_1,\dots,X_n)^{\mathrm{T}}$ takes a vector $\boldsymbol{x}\in\mathcal{X}^{\boldsymbol{n}}$ with probability $p_{\boldsymbol{x}}$.

周辺確率分布(Marginal probability distribution)

$\boldsymbol{X}=\left(\boldsymbol{X}_1^{\mathrm{T}},\boldsymbol{X}_2^{\mathrm{T}}\right)^{\mathrm{T}}$と$\boldsymbol{x}=\left(\boldsymbol{x}_1^{\mathrm{T}},\boldsymbol{x}_2^{\mathrm{T}}\right)^{\mathrm{T}}\in\mathcal{X}^{n_1+n_2}\,(n_1+n_2=n)$に対して、

For $\boldsymbol{X}=\left(\boldsymbol{X}_1^{\mathrm{T}},\boldsymbol{X}_2^{\mathrm{T}}\right)^{\mathrm{T}}$ and $\boldsymbol{x}=\left(\boldsymbol{x}_1^{\mathrm{T}},\boldsymbol{x}_2^{\mathrm{T}}\right)^{\mathrm{T}}\in\mathcal{X}^{n_1+n_2}$ with $n_1+n_2=n$,

$$P(\boldsymbol{X}_1=\boldsymbol{x}_1)=\sum_{\boldsymbol{x}_2\in\mathcal{X}^{n_2}}P(\boldsymbol{X}_1=\boldsymbol{x}_1,\boldsymbol{X}_2=\boldsymbol{x}_2)$$

条件付き確率分布(Conditional probability distribution)

$$P(\boldsymbol{X}_2=\boldsymbol{x}_2|\,\boldsymbol{X}_1=\boldsymbol{x}_1)=\frac{P(\boldsymbol{X}_1=\boldsymbol{x}_1,\boldsymbol{X}_2=\boldsymbol{x}_2)}{P(\boldsymbol{X}_1=\boldsymbol{x}_1)}.$$

独立性(Independence)

$$P(\boldsymbol{X}=\boldsymbol{x})=\prod_{i=1}^{n}P(X_i=x_i).$$

n次元連続確率ベクトル(n-dimensional continuous random vector)

同時確率密度関数(Joint pdf)

$$P(\boldsymbol{X} \in A) = \int_A p_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}.$$

連続確率ベクトル$\boldsymbol{X} = (X_1, \dots, X_n)^{\mathrm{T}}$が集合$A \subset \mathbb{R}^n$に入る確率

Probability with which a continuous random vector $\boldsymbol{X} = (X_1, \dots, X_n)^{\mathrm{T}}$ is in a subset $A \in \mathbb{R}^n$.

周辺確率密度関数(Marginal pdf)

$\boldsymbol{X} = \left(\boldsymbol{X}_1^{\mathrm{T}}, \boldsymbol{X}_2^{\mathrm{T}}\right)^{\mathrm{T}}$と$\boldsymbol{x} = \left(\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}\right)^{\mathrm{T}} \in \mathcal{X}^{n_1+n_2}$ $(n_1 + n_2 = n)$に対して、

For $\boldsymbol{X} = \left(\boldsymbol{X}_1^{\mathrm{T}}, \boldsymbol{X}_2^{\mathrm{T}}\right)^{\mathrm{T}}$ and $\boldsymbol{x} = \left(\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}\right)^{\mathrm{T}} \in \mathcal{X}^{n_1+n_2}$ with $n_1 + n_2 = n$,

$$p_{\boldsymbol{X}_1}(\boldsymbol{x}_1) = \int_{\mathbb{R}^{n_2}} p_{\boldsymbol{X}}(\boldsymbol{x}_1, \boldsymbol{x}_2) d\boldsymbol{x}_2.$$

条件付き確率密度関数(Conditional pdf)

$$p_{\boldsymbol{X}_1|\boldsymbol{X}_2}(\boldsymbol{x}_1|\boldsymbol{x}_2) = \frac{p_{\boldsymbol{X}}(\boldsymbol{x}_1, \boldsymbol{x}_2)}{p_{\boldsymbol{X}_2}(\boldsymbol{x}_2)}.$$

独立性(Independence)

$$p_{\boldsymbol{X}}(\boldsymbol{x}) = \prod_{i=1}^{n} p_{X_i}(x_i).$$

共分散行列(Covariance matrix)

確率ベクトル$\boldsymbol{X} = (X_1, \dots, X_n)^{\mathrm{T}}$に対して、$n \times n$行列$\boldsymbol{\Sigma}$の$(i, j)$要素$[\boldsymbol{\Sigma}]_{i,j}$が相関$\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])]$に等しいとき、$\boldsymbol{\Sigma}$は共分散行列と呼ばれる。

An $n \times n$ matrix $\boldsymbol{\Sigma}$ is called covariance matrix if the (i, j)th element $[\boldsymbol{\Sigma}]_{i,j}$ is given by $\mathbb{E}[(X_i -$

性質1(Property 1)　共分散行列は対称行列である。(Any covariance matrix is symmetric.)

$$\because \ [\boldsymbol{\Sigma}]_{j,i} = \mathbb{E}\left[\left(X_j - \mathbb{E}[X_j]\right)(X_i - \mathbb{E}[X_i])\right] = \mathbb{E}\left[(X_i - \mathbb{E}[X_i])\left(X_j - \mathbb{E}[X_j]\right)\right] = [\boldsymbol{\Sigma}]_{i,j}.$$

性質2(Property 2)　共分散行列は半正定値である。すなわち、

Any covariance matrix is positive semi-definite, i.e.

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a} \geq 0 \quad \text{for all } \boldsymbol{a} \in \mathbb{R}^n.$$

$\because$　共分散行列は、$\boldsymbol{\Sigma} = \mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^{\mathrm{T}}]$と表現できる。

The covariance matrix can be represented as $\boldsymbol{\Sigma} = \mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^{\mathrm{T}}]$.

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{a} = \boldsymbol{a}^{\mathrm{T}}\mathbb{E}[(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^{\mathrm{T}}]\boldsymbol{a} = \mathbb{E}[\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^{\mathrm{T}}\boldsymbol{a}]$$

$$= \mathbb{E}[\{(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^{\mathrm{T}}\boldsymbol{a}\}^{\mathrm{T}}(\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^{\mathrm{T}}\boldsymbol{a}] = \mathbb{E}\left[\left((\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}])^{\mathrm{T}}\boldsymbol{a}\right)^2\right] \geq 0.$$

期待値の性質(Properties of Expectation)

性質１(Property 1)　任意の確率変数XとYに対して、

For all random variables X and Y,

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y].$$

性質２(Property 2)　独立な確率変数列$\{X_i\}_{i=1}^{n}$に対して、

For independent random variables $\{X_i\}_{i=1}^{n}$,

$$\mathbb{E}\left[\prod_{i=1}^{n} g_i(X_i)\right] = \prod_{i=1}^{n} \mathbb{E}[g_i(X_i)].$$

性質３(Property 3)

$$\mathbb{E}[f(X)g(Y)|Y = y] = g(y)\mathbb{E}[f(X)|Y = y].$$

無相関な確率変数(Uncorrelated random variables)

二つの確率変数XとYは、以下の条件を満たすとき、無相関と呼ばれる。

Two random variables X and Y are called uncorrelated if the following holds:

$$\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]=0.$$

性質4(Property 4) $\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]=\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y].$

$$\because \quad \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]=\mathbb{E}[XY-X\mathbb{E}[Y]-\mathbb{E}[X]Y+\mathbb{E}[X]\mathbb{E}[Y]]$$

$$=\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]-\mathbb{E}[X]\mathbb{E}[Y]+\mathbb{E}[X]\mathbb{E}[Y]=\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y].$$

性質5(Property 5)

二つの確率変数XとYが独立ならば、XとYは無相関である。

If two random variables X and Y are independent, they are uncorrelated.

$\because$ 期待値の性質2から、(From Property 2 of expectation, we have)

$$\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]=\mathbb{E}[X-\mathbb{E}[X]]\mathbb{E}[Y-\mathbb{E}[Y]]$$

$$=(\mathbb{E}[X]-\mathbb{E}[X])(\mathbb{E}[Y]-\mathbb{E}[Y])=0.$$

注意(Remark): 逆は正しくない。(The converse is not correct.)

分散の和(Sum of variance)

定理1.1(Theorem 1.1)

$\{X_i\}_{i=1}^n$を無相関な確率変数列とする。和$Y = \sum_{i=1}^n X_i$の分散は、分散の和に等しい。

Let $\{X_i\}_{i=1}^n$ denote uncorrelated random variables. The variance of the sum $Y = \sum_{i=1}^n X_i$ is equal to the sum of the variance.

$$\mathbb{V}[Y] = \sum_{i=1}^n \mathbb{V}[X_i].$$

注意(Remark)

無相関性をより強い条件である独立性に置き換えても、当然、定理は従う。

The theorem is of course correct, when the assumption of uncorrelated variables is replaced with independence, which is a stronger condition.

定理1.1の証明(Proof of Theorem 1.1)

$\tilde{X}_i = X_i - \mathbb{E}[X_i]$とおくことで、一般性を失うことなく、$\mathbb{E}[X_i] = 0$を仮定できる。

Without loss of generality, we can assumed $\mathbb{E}[X_i] = 0$ by letting $\tilde{X}_i = X_i - \mathbb{E}[X_i]$.

期待値の性質1から、$\mathbb{E}[Y] = \sum_i \mathbb{E}[X_i] = 0$が従う。それゆえ、

Property 1 of expectation implies $\mathbb{E}[Y] = \sum_i \mathbb{E}[X_i] = 0$. Thus, we have

$$\mathbb{V}[Y] = \mathbb{E}[Y^2] = \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}[X_i X_j],$$

最後の等号は、期待値の性質1のためである。

where the last equality is due to Property 1 of expectation.

無相関仮定$\mathbb{E}[X_i X_j] = 0 \ (i \neq j)$を使うために、総和を二つに分ける。

We decompose the summation into two terms to utilize the assumption of uncorrelated variables,

$$\mathbb{V}[Y] = \sum_{i=1}^{n} \mathbb{E}[X_i^2] + \sum_{i \neq j} \mathbb{E}[X_i X_j] = \sum_{i=1}^{n} \mathbb{E}[X_i^2] = \sum_{i=1}^{n} \mathbb{V}[X_i],$$

最後の等号は、$\mathbb{E}[X_i] = 0$から従う。

where the last equality follows from $\mathbb{E}[X_i] = 0$. ■