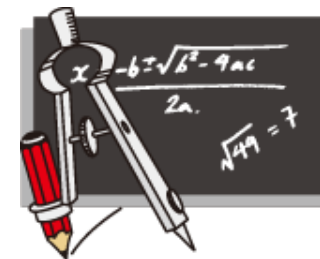


応用数学III

(4)多次元確率分布

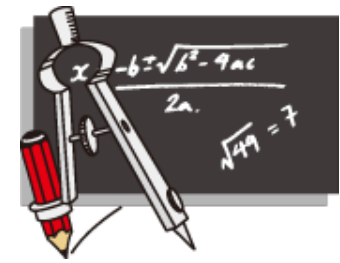
木村真一

講義のスケジュール

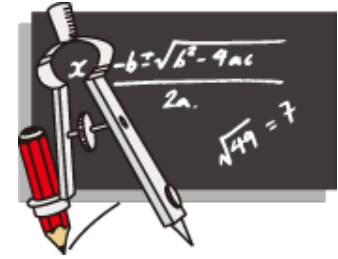


- (1) 確率の基礎
- (2) 確率変数と確率分布
- (3) いろいろな確率分布
- (4) 多次元確率分布**
- (5) 大数の法則と中心極限定理
- (6) 確率過程の基礎 1
- (7) 確率過程の基礎 2
- (8) フーリエ解析
- (9) フーリエ変換の性質
- (10) 相関解析
- (11) 群・環・体の定義
- (12) 準同型写像
- (13) N を法とする合同式
- (14) 線形代数

参考書 1



参考書 2



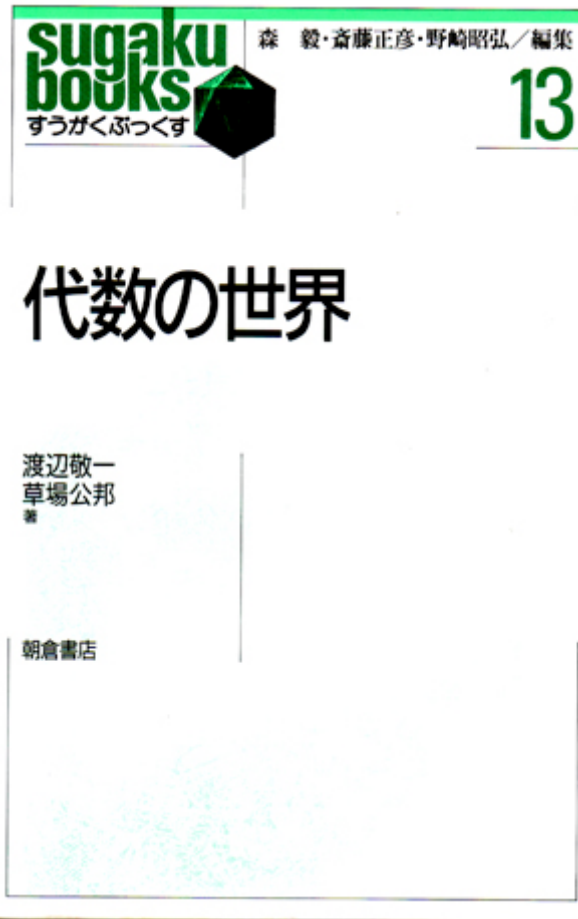
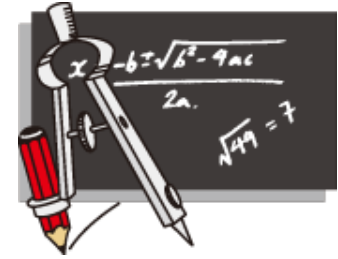
信号解析入門

越川常治 著



近代科学社

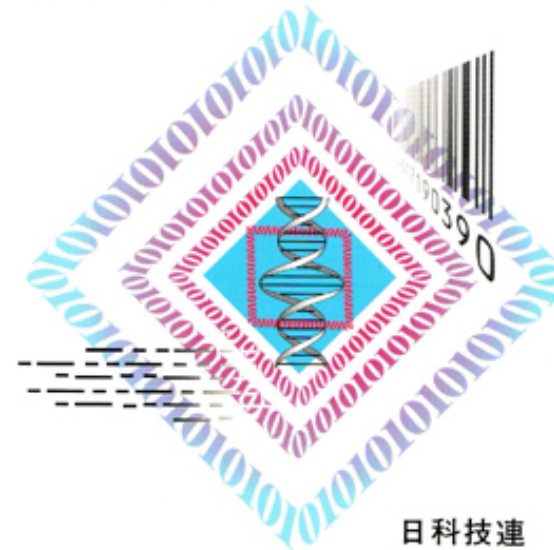
参考書 3



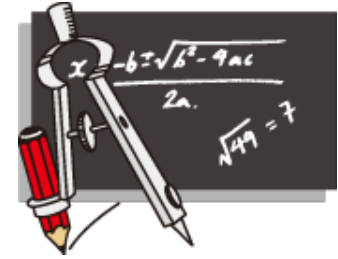
情報数学のはなし

●情報理論から暗号・認証まで

大村 平 著



前回の復習：確率分布



- 一様分布

$$f(x) = \frac{1}{b-a}, X \sim U(a, b)$$

- 二項分布

$$f(x) = P(X = x) = \frac{n!}{(n-x)!x!} \theta^x (1-\theta)^{n-x}, X \sim B(n; \theta)$$

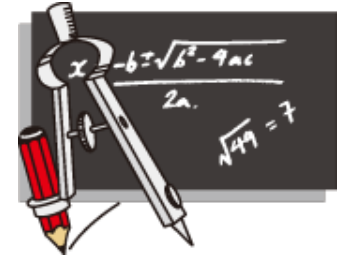
- ポアソン分布

$$f(x) = P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, X \sim Po(\lambda)$$

- 指数分布

$$f(x) = \lambda \cdot e^{-\lambda x}, X \sim Ex(\lambda)$$

訂正：二項分布の平均



$$\begin{aligned}\sum xf(x) &= \sum_{x=0}^n \frac{n!}{(n-x)!x!} x \theta^x (1-\theta)^{n-x} \\ &= \sum_{x=1}^n \frac{n!}{(n-x)!(x-1)!} \theta^x (1-\theta)^{n-x} \\ &= \theta \sum_{x=1}^n \frac{n!}{(n-x)!(x-1)!} \theta^{(x-1)} (1-\theta)^{n-x}\end{aligned}$$

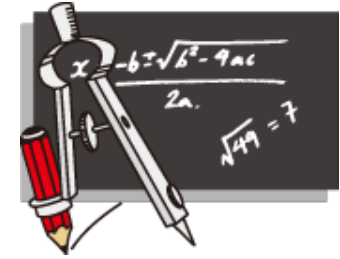
x-1=kで変数変換

$$= \theta \sum_{k=0}^{n-1} \frac{n!}{(n-k-1)!k!} \theta^k (1-\theta)^{n-k-1}$$

m=n-1で変数変換

$$\begin{aligned}&= \theta \sum_{k=0}^m \frac{(m+1)!}{(m-k)!k!} \theta^k (1-\theta)^{m-k} \\ &= \theta(m+1) \sum_{k=0}^m \frac{m!}{(m-k)!k!} \theta^k (1-\theta)^{m-k} \\ &= n\theta \{\theta + (1-\theta)\}^m = n\theta\end{aligned}$$

訂正：二項分布の分散 1

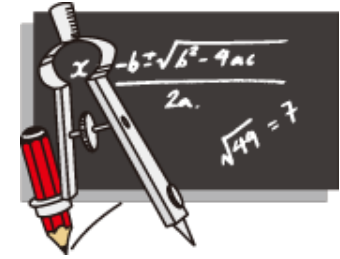


$$\begin{aligned} & \sum x^2 f(x) \\ &= \sum_{x=0}^n \frac{n!}{(n-x)!x!} x^2 \theta^x (1-\theta)^{n-x} \\ &= \theta \sum_{x=1}^n \frac{n!}{(n-x)!(x-1)!} x \theta^{x-1} (1-\theta)^{n-x} \end{aligned}$$

$x-1=k$ 、 $m=n-1$ で変数変換

$$\begin{aligned} &= \theta(m+1) \sum_{k=0}^m \frac{m!}{(m-k)!k!} (k+1) \theta^k (1-\theta)^{m-k} \\ &= n\theta \sum_{k=0}^m \frac{m!}{(m-k)!k!} (k+1) \theta^k (1-\theta)^{m-k} \\ &= n\theta \left\{ \sum_{k=1}^m \frac{m!}{(m-k)!k!} k \theta^k (1-\theta)^{m-k} + \sum_{k=0}^m \frac{m!}{(m-k)!k!} \theta^k (1-\theta)^{m-k} \right\} \end{aligned}$$

訂正：二項分布の分散 2



$$= n\theta \sum_{k=1}^m \frac{m!}{(m-k)!k!} k \theta^k (1-\theta)^{m-k} + n\theta$$

$$= n\theta \cdot m\theta \sum_{k=1}^m \frac{(m-1)!}{(m-k)!(k-1)!} \theta^{k-1} (1-\theta)^{m-k} + n\theta$$

$$\begin{aligned} \sum (x - \bar{x})^2 f(x) &= n\theta(n-1)\theta + n\theta - \bar{x}^2 \\ &= n^2\theta^2 - n\theta^2 + n\theta - n^2\theta^2 \\ &= n\theta(1-\theta) \end{aligned}$$

再び $k-1=i$ 、 $j=m-1$ で変数変換

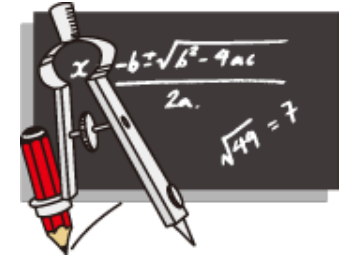
$$= n\theta \cdot m\theta \sum_{i=0}^j \frac{j!}{(j-i)!i!} \theta^i (1-\theta)^{j-i} + n\theta$$

$$= n\theta \cdot m\theta \cdot (\theta + 1 - \theta)^j + n\theta$$

$$= n\theta m\theta + n\theta$$

$$= n\theta(n-1)\theta + n\theta = n^2\theta^2 - n\theta^2 + n\theta$$

指数分布の平均と分散



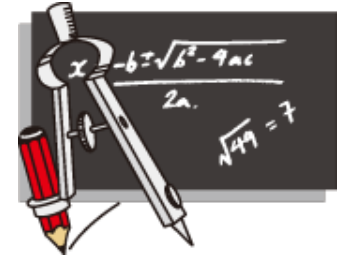
- 指数関数の平均と分散を計算してみてください。

- 平均 $\frac{1}{\lambda}$

- 分散 $\frac{1}{\lambda^2}$

- 応用例としては、災害が起こる時間間隔、機械が突発的な理由で故障する時間間隔などがあります。

正規分布



- 小生の主義に大変反するのですが、ここから先は説明が天下りになります。（理由があります。）
- 次のような関数を考えます。

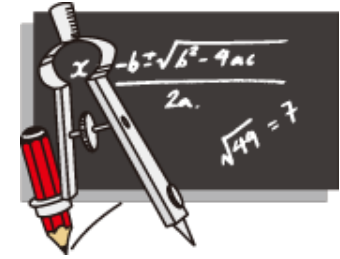
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, X \sim N(\mu, \sigma^2)$$

- 特徴としては
 - $-\infty$ から ∞ まで広がっているけど、ほとんどは平均値付近に集まっている（ 3σ 内に99.73%）
 - 平均値に対して左右対称

など観測を行ったときのデータのばらつきによく似た性質を持っています。

- この分布を正規分布といいます。

正規分布 (つづき)

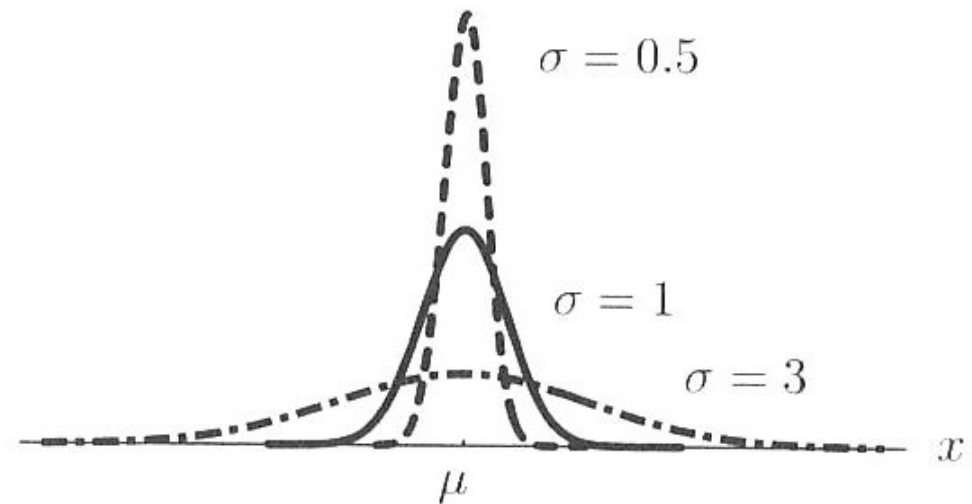


- 平均と分散はどのようになるでしょう。

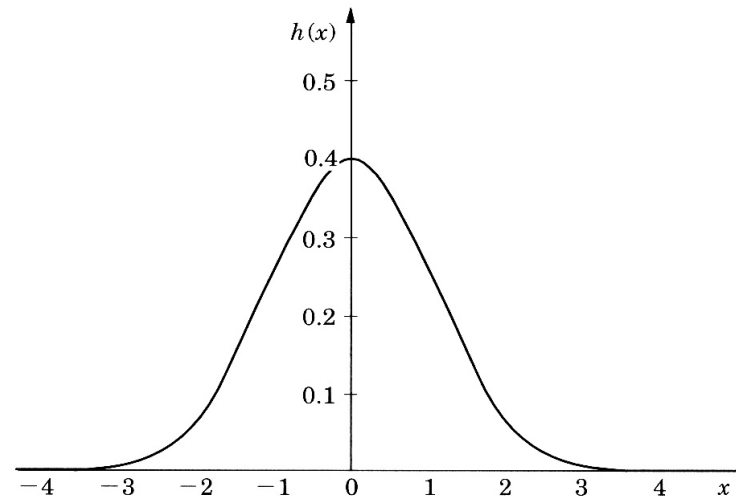
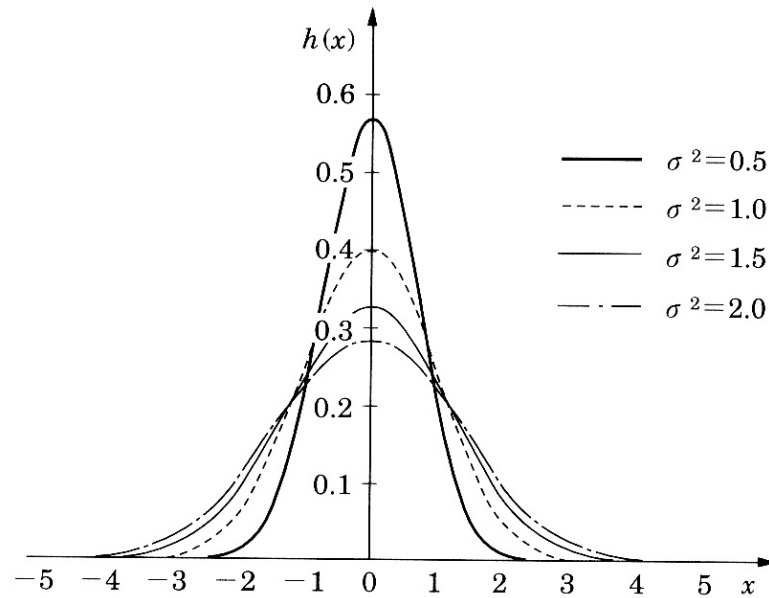
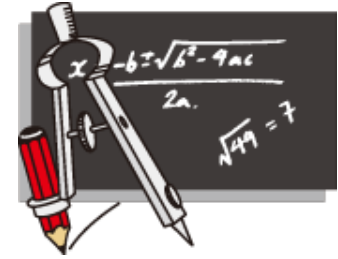
- 平均は
 μ になります。

- 分散は
 σ^2 になります。

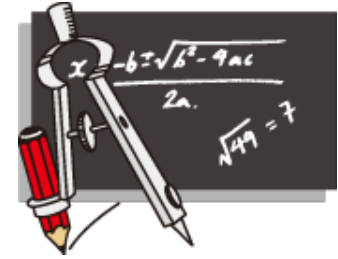
- 実はこの正規分布、ある範囲を設定したときの確率も解析的には計算できません！！
- ですので、数表か計算機による計算で代用します。



正規分布



多次元確率変数



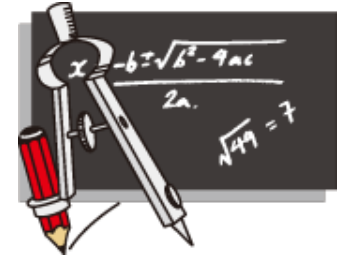
- これまでは確率変数は1次元でした。
- 複数の変数を同時に取り扱いたい事がよくあります。
 - 例えば身長と体重の分布、先の袋から玉を取り出すくじで玉の色と玉に描かれた数字等々
- そこで確率変数を組み合わせてベクトルとして扱います。

$$\mathbf{X} = (X_1, X_2, \dots, X_n)^t$$

- これを同時確率分布といいます。
- 分布関数は次のように表されます。

$$\begin{aligned} P(\mathbf{X} \leq \mathbf{x}) &= P(X_1 \leq x_1, \dots, X_n \leq x_n) \\ &= \int_{-\infty}^{x_n} \dots \int_{-\infty}^{x_1} f(x_1, \dots, x_n) dx_1 \dots dx_n \end{aligned}$$

同時確率分布と同時確率密度分布



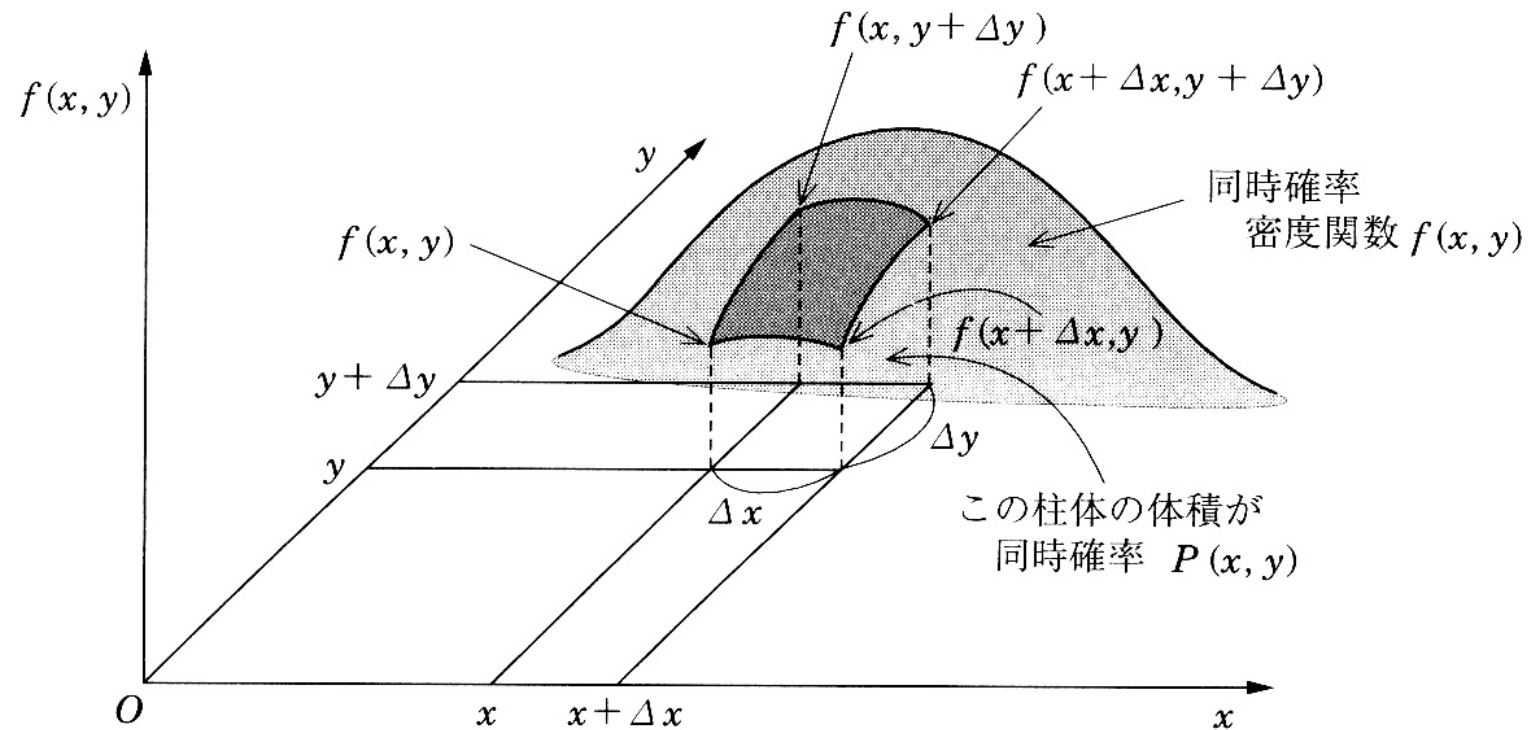
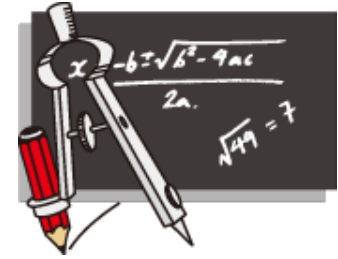
- 分かりやすい例として2次元を考えましょう。
- 下の図のように X と Y 二つの確率変数があるとすると、同時確率分布関数と同時確率密度分布関数は次のように定義されます。
- 同時確率密度関数

$$P(x \leq X \leq x + \Delta x, y \leq Y \leq y + \Delta y)$$
$$= \int_x^{x+\Delta x} \int_y^{y+\Delta y} f(x', y') dy' dx'$$

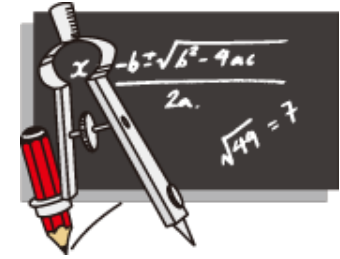
- 同時確率分布関数

$$F(x, y) = P(X \leq x, Y \leq y)$$
$$= \int_{-\infty}^x \int_{-\infty}^y f(x', y') dy' dx'$$

同時確率分布



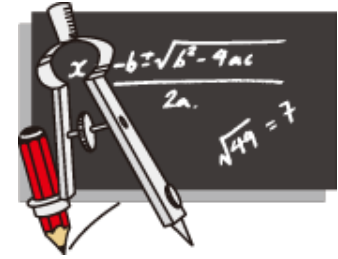
周辺分布関数



- これに対し、多次元確率分布変数の1つの分布だけに着目した分布関数のことを、周辺分布関数といいます。
- 次のように定義されます。

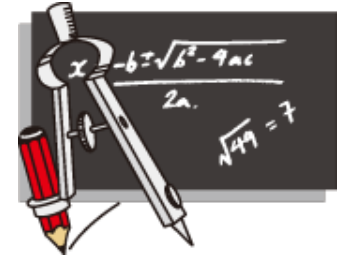
$$f_1(x) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_2 \dots dx_n$$

演習問題 1



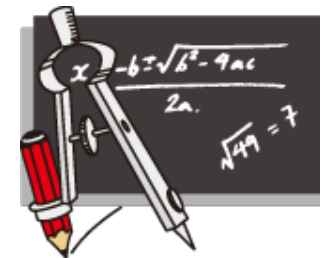
- 2次元確率密度関数 $f(x,y)$ が、ある矩形領域($x_1 \leq X \leq x_2$, $y_1 \leq Y \leq y_2$)で一定値をとるとき、この分布は2次元一様分布といわれる。このときの一定値を求めよ。

演習問題 2



- ある地点での気温と湿度を時刻を変えて同時に計測する場合、2次元同時確率分布、及びその周辺分布はどのようにしたら計測できるか。

多次元確率変数の期待値



- 多次元確率変数においても期待値は同様に定義されます。但し結果はベクトルです。

$$E[g(\mathbf{X})] = \int g(\mathbf{x})f(\mathbf{x})d\mathbf{x}$$

$$E[g(\mathbf{X})] = (E[g_1(\mathbf{x})], \dots, E[g_n(\mathbf{x})])$$

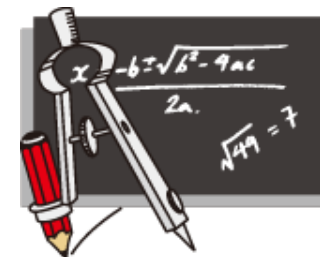
- 平均値も平均ベクトルとして定義されます。

$$\vec{\mu} = (\mu_1, \dots, \mu_n) = E[\mathbf{X}] = (E[X_1], \dots, E[X_n])$$

- 以下のように、各要素は周辺確率分布の平均値と同じになります。

$$\begin{aligned} E[X_1] &= \int x_1 f(\mathbf{x}) d\mathbf{x} = \int x_1 \int \dots \int f(\mathbf{x}) dx_n \dots dx_1 \\ &= \int x_1 f(x_1) dx_1 \end{aligned}$$

共分散



- 次に多次元確率変数の分散を求めてみましょう。理屈は1次元と同じです。

$$\sigma_i^2 = V[X_i] = E[(X_i - \mu_i)^2]$$

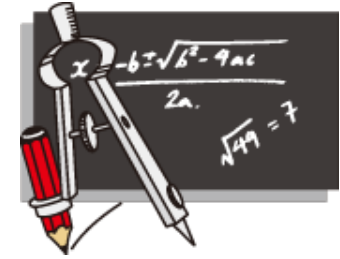
- ただ、ここで単に偏差の自乗だけでなく、偏差を他の要素と組み合わせると、変数の間の関係がわかってきて便利です。
- そこで次のような量を共分散として定義します。

$$\sigma_{ij} = Cov[X_i, X_j] = E[(X_i - \mu_i)(X_j - \mu_j)]$$

- このような共分散をまとめて共分散行列として取り扱います。半正定値対角行列

$$\Sigma = V[\mathbf{X}] = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}$$

相関係数

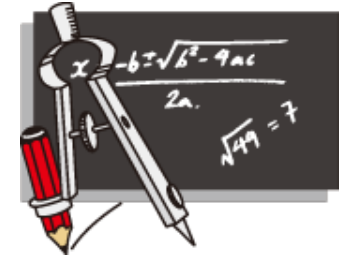


- 共分散の他に確率変数の間の関係を表す指標として相関係数があります。

$$\rho_{ij} = \text{Corr}[X_i, X_j] = \frac{\text{Cov}[X_i, X_j]}{\sqrt{V[X_i]V[X_j]}}$$

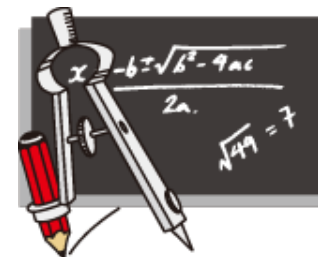
- 標準化した変数の共分散です。
- 相関係数が0の時、二つの確率変数の間に相関はありません。
- 相関係数が1の時、二つの変数の間に線形性が成り立ちます。

演習問題 3



- 1,1,2,3と書かれた4枚のカードから2枚抜き出して並べるとき、はじめのカードの数を X 、2番目のカードの数を Y とする。
 - X, Y の同時分布を求め、その周辺分布を求めよ。
 - X, Y それぞれの平均、及び分散を求めよ。
 - X, Y の共分散を求めよ。

確率変数の独立性



- 複数の事象が独立であるとき、ある事象の組み合わせの確率は、それぞれの事象の積になりました。

$$P((X_1, \dots, X_n) = (x_1, \dots, x_n)) = P(X_1 = x_1) \cdot \dots \cdot P(X_n = x_n)$$

- 事象をもとにした確率では独立性を以下の様に表しました。

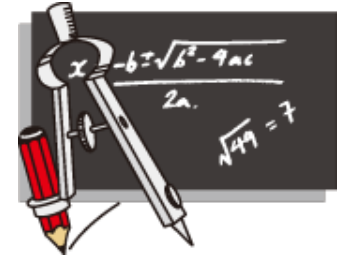
$$P(A \cap B) = P(A)P(B)$$

- これを連続型の確率変数に応用します。確率変数が以下の条件を満たすとき、独立であるといいます。

$$f(x_1, \dots, x_n) = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

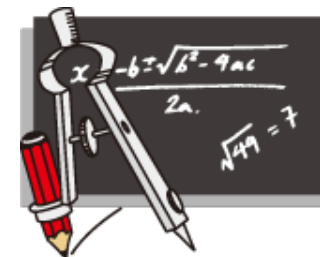
$$F(x_1, \dots, x_n) = F_1(x_1) \cdot \dots \cdot F_n(x_n)$$

確率変数の独立性



- 表現を変えれば、同時確率分布が周辺確率分布の積で表されるとき、各確率変数は独立であるといえます。
- 特に独立が確率変数 X_1, \dots, X_n が同一の確率分布 $F(x)$ に従っているとき確率変数 X_1, \dots, X_n は、独立同一分布に従うといえます。

確率変数の独立性と期待値



- 確率変数が独立であるならば次の関係が成り立ちます。

$$E[X_1 \cdot \dots \cdot X_n] = E[X_1] \cdot \dots \cdot E[X_n]$$

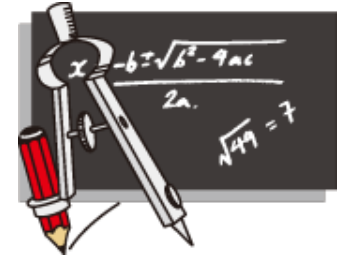
- 例として $n=2$ の時を示します。

$$\begin{aligned} E[X_1 \cdot X_2] &= \int \int x_1 x_2 f(x_1, x_2) dx_1 dx_2 = \int \int x_1 x_2 f_1(x_1) f_2(x_2) dx_1 dx_2 \\ &= \int x_1 f_1(x_1) dx_1 \cdot \int x_2 f_2(x_2) dx_2 = E[X_1] \cdot E[X_2] \end{aligned}$$

- また、独立であるならば二つの確率変数は無相関です。

$$\begin{aligned} Cov[X, Y] &= E[(X - \mu_X) \cdot (Y - \mu_Y)] = E[(X - \mu_X)] \cdot E[(Y - \mu_Y)] \\ &= 0 \end{aligned}$$

確率変数の和の平均（例題）

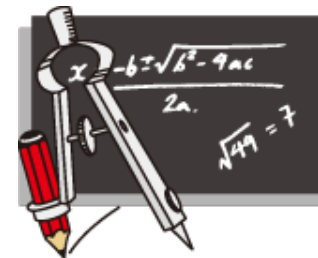


- 確率変数の和の平均はどのようになりますか。
- 今までの例を参考に考えてみてください。

$$\begin{aligned} E[X_1 + X_2] &= \iint (x_1 + x_2) f(\mathbf{x}) dx_1 dx_2 \\ &= \iint x_1 f(\mathbf{x}) dx_1 dx_2 + \iint x_2 f(\mathbf{x}) dx_1 dx_2 \\ &= E[X_1] + E[X_2] \end{aligned}$$

- おっとこれは簡単、各平均の和です。

確率変数の和の分散（例題）

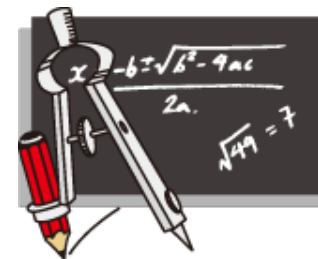


- それでは分散は？

$$\begin{aligned} V[X_1 + X_2] &= E\left[\left((X_1 + X_2) - (\mu_1 + \mu_2)\right)^2\right] \\ &= E\left[\left((X_1 - \mu_1) + (X_2 - \mu_2)\right)^2\right] \\ &= E\left[(X_1 - \mu_1)^2 + 2(X_1 - \mu_1) \cdot (X_2 - \mu_2) + (X_2 - \mu_2)^2\right] \\ &= E\left[(X_1 - \mu_1)^2\right] + E\left[(X_2 - \mu_2)^2\right] + 2E\left[(X_1 - \mu_1) \cdot (X_2 - \mu_2)\right] \\ &= V[X_1] + V[X_2] + 2Cov[X_1, X_2] \end{aligned}$$

- 単純に各分散の和にはならず共分散が入り込みます。
- ですから、それぞれ独立ならば、各分散の和になります。

条件付き確率変数



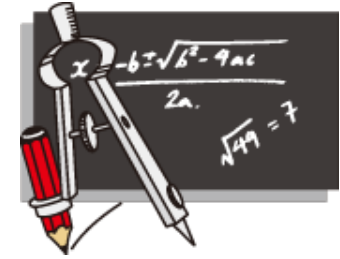
- 周辺確率分布を用いると、事象をもとにした確率で議論した条件付き確率を、確率変数に拡張できます。
- 事象をもとにした確率では条件付き確率は次のように表されました。

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

- 確率変数では以下のように表します。

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)}$$

チェビシェフの不等式

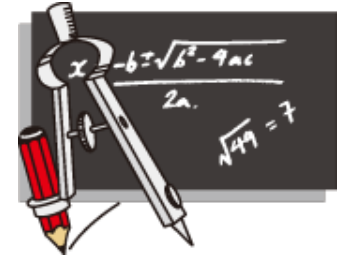


- 確率変数 X の平均値が μ 、分散が σ であるとき、任意の $\varepsilon > 0$ について、次の不等式が成り立ちます

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

- 分散 σ が小さいほど平均付近に集中している

コーシー・シュバルツの不等式

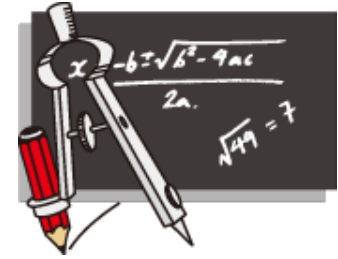


- 確率変数 X と Y があるとき、分散と共分散の間には次の不等式が成り立ちます。

$$\text{Cov}[X, Y]^2 \leq V[X] \cdot V[Y]$$

- 相関係数は $\rho = \frac{\text{Cov}[X, Y]}{\sqrt{V[X]V[Y]}}$ でしたから、絶対値は必ず1以下です。

まとめ



- 正規分布
- 同時確率分布
- 周辺分布関数
- 共分散
- 相関係数
- 確率変数の独立性
- 条件付確率変数
- チェビシェフの不等式
- コーシー・シュバルツの不等式