情報理論

2017年4月14日, 4月21日

羽石　秀昭

スケジュール

1. 情報理論の概要・情報の表現　4/14
2. 確率の基礎　4/21
3. 情報量（エントロピー、ダイバージェンス、相互情報量） 4/28
4. 情報量の性質　5/2　(5/4祝日）
5. 演習　5/11　（羽石不在）
6. 情報源のモデルとエントロピーレート 5/19
7. 情報源の符号化　5/26
8. 中間テスト 6/2　（羽石不在）
9. 相互情報量（１）基礎　6/9
10. 相互情報量（２）応用　6/16
11. 演習　6/23　（羽石不在？）
12. 情報量統計学（１）最尤法 6/30
13. 情報量統計学（２）ＡＩＣ 7/7
14. 情報理論の応用（１）ベイズ推定 7/14
15. 情報理論の応用（２）パターン認識　7/21
16. 期末テスト 7/28

情報の表現

集合の表現

アルファベットと符号化

集合の表現

1．集合の表現方法

要素を{}で括る。

例）　$B=\{0,\ 1\}$

$A=\{x: x\text{ の満たす条件}\}$あるいは$A=\{x|x\text{ の満たす条件}\}$などと表す。

例）　$P=\{n : n\text{ は素数}\}$

2．和集合、積集合

例）　$A=\{1,2,3\},\ B=\{2,4,5\}$ のとき,

A と B の和集合は　$A\cup B=\{1,2,3\}\cup\{2,4,5\}=\{1,2,3,4,5\}$

A と B の積集合は　$A\cup B=\{1,2,3\}\cap\{2,4,5\}=\{2\}$

3．部分集合、補集合

例）　集合Ｂは集合Ａの部分集合である：　$A\supset B$

例）集合Ａの補集合　$\bar{A}$

4．要素と集合の関係

例）集合Aはaを要素にもつ:　　$a\in A$

集合の表現：直積

順序対

２つの要素 x, y をこの順番で並べた組(x, y) のこと。
順序対では並べる順序が違っている組は違ったものとして扱うので,
$x \neq y$ のとき, $(x, y) \neq (y, x)$

直積

集合 A と B からそれぞれ１つずつ要素を取ってきて作った順序対のすべてを要素とする集合を

$$A \times B = \{(a, b): a \in A, b \in B\}$$

によって表し, A と B の直積と呼ぶ。

例）

$A = \{a, b\}$, $B = \{1,2,3\}$ とするとき,

$$A \times B = \{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3)\}$$
$$B \times A = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$$
$$A \times A = \{(a, a), (a, b), (b, a), (b, b)\}$$

であり, この場合, $A \times B \neq B \times A$ であることがわかる

アルファベットと符号化

情報理論では取り扱うデータを表す記号の集合のことをアルファベットと呼ぶ。
たとえば，数字のアルファベットは　A={0,1,2,3,4,5,6,7,8,9}　となる.
一般にアルファベットA がA=$\{a_1, a_2, \dots, a_N\}$

と書かれたとき，アルファベット A に含まれる要素のことを記号といい，アルファベットA に含まれる記号の総数N をアルファベットのサイズあるいは大きさという.

アルファベットA に対し, A の記号をk 個並べてできる長さk の記号列の集合をA^k によって表す

長さ0 以上のA 上の記号列全体の集合を

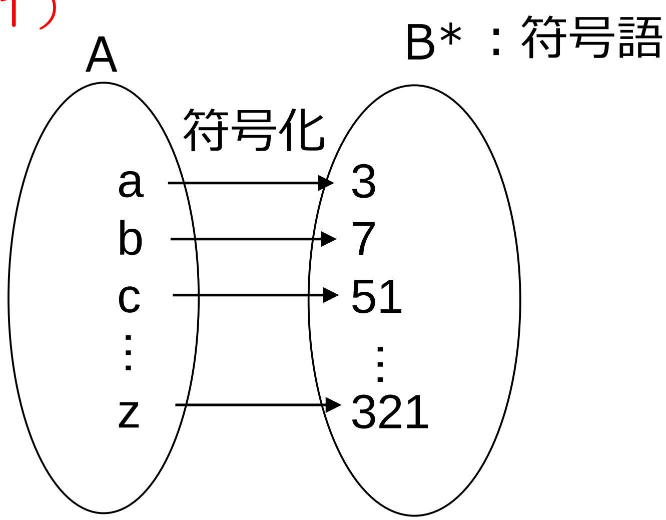
A^*

によってあらわす。

アルファベットと符号化

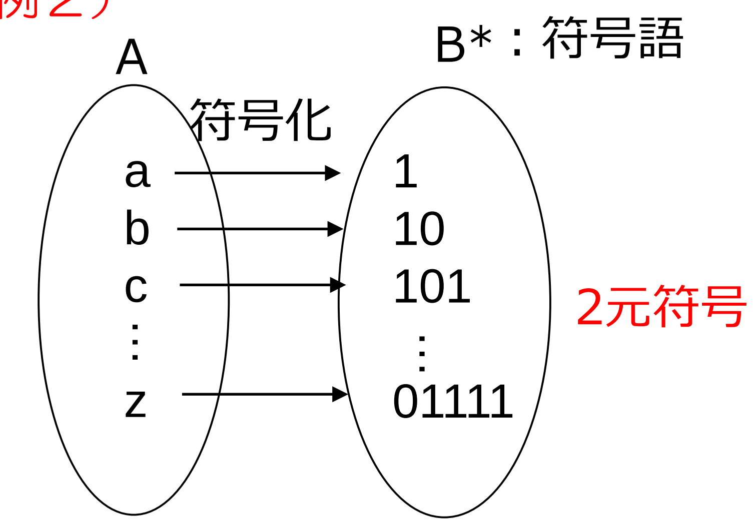
2つのアルファベットA とB があり, A の記号にB 上の記号列, すなわちB*の記号列を対応させる写像のことを符号化という。

例１）



$A=\{a,b,\ldots,z\}$　$B=\{0,1,\ldots,9\}$

例２）



$A=\{a,b,\ldots,z\}$　$B=\{0,1\}$

このとき, B を符号アルファベットとも呼ぶ.

特に符号アルファベットがB ={0, 1}である符号を2元符号と呼ぶ.

符号語の長さがA の記号によらず一定の場合を

固定長符号, そうでない場合を可変長符号と呼ぶ.

ASCII code （符号）：固定長符号

		上位3ビット							
		0	1	2	3	4	5	6	7
下位4ビット	0	（null文字）	（データリンクエスケープ）	（スペース）	0	@	P	`	p
	1	（ヘッダ開始）	（装置制御1）	!	1	A	Q	a	q
	2	（テキスト開始）	（装置制御2）	″	2	B	R	b	r
	3	（テキスト終了）	（装置制御3）	#	3	C	S	c	s
	4	（転送終了）	（装置制御4）	$	4	D	T	d	t
	5	（照会）	（受信失敗）	%	5	E	U	e	u
	6	（受信OK）	（同期）	&	6	F	V	f	v
	7	（警告）	（転送ブロック終了）		7	G	W	g	w
	8	（後退）	（取り消し）	(	8	H	X	h	x
	9	（水平タブ）	（メディア終了）	)	9	I	Y	i	y
	A	（改行）	（置換）	*	:	J	Z	j	z
	B	（垂直タブ）	（エスケープ）	+	;	K	[	k	{
	C	（改ページ）	（フォーム区切り）	,	<	L	¥	l	|
	D	（復帰）	（グループ区切り）	−	=	M	]	m	}
	E	（シフトアウト）	（レコード区切り）	.	>	N	^	n	~
	F	（シフトイン）	（ユニット区切り）	/	?	O	_	o	（削除）

http://marunouchi-tech.i-studio.co.jp/1973/

確率の基礎

- ✓ 事象と確率
- ✓ 確率の定義
- ✓ 条件付き確率と事象の独立性
- ✓ 確率変数と確率分布
- ✓ 平均
- ✓ 分散

事象と確率

標本空間：サイコロを振ったとき出る目が1から6までであるように, 偶然を伴う実験の起こり得る結果全体の集合を標本空間と呼び, Ω（オメガ）によって表す.

例１）サイコロの目を振ったとき出る目の標本空間は, Ω＝｛1, 2, 3, 4, 5, 6｝
例２）コイン投げの結果を表す標本空間は, Ω = { 表, 裏 }

事象：標本空間の部分集合のことを事象という。

実験の結果, 事象 E の要素の1つが観測されたとき「事象 E が生じた」 あるいは「事象E が起きた」という.

サイコロの例で, 事象 E を $E = \{2, 4, 6\}$ によって定めれば, サイコロを振って偶数の目が出たときに, 事象 E が生じたことになる.

根元事象：ただ1つの要素からなる事象のこと

空集合：どの結果も含まない事象。 記号$\emptyset$によって表す.

余事象：事象E に属さない根元事象全体の集まり, すなわち事象E の標本空間 Ω に対する補集合をE の余事象といい, $\bar{E}$によって表す。

確率の定義

標本空間Ω の事象Eに対して、次の３条件を満たすように
$P(E)$を対応させることができるとき、 Pを確率という。

１．任意の事象Eに対して

$$0 \leq P(E) \leq 1$$

２．標本空間Ωに対して

$$P(\Omega) = 1$$

３．事象が互いに排反なら,すなわち$\mathrm{E}_i \cup E_j = \emptyset \quad (\mathrm{i} \neq j)$が成り立つとき

$$P(\mathrm{E}_1 \cup E_2 \cup \cdots) = P(\mathrm{E}_1) + P(E_2) + \cdots$$

条件付き確率と事象の独立性

条件付き確率：

２つの事象Ｅ，Ｅ′があり、事象Ｅ′の生じる確率がP(E′)>0ならば、事象Ｅ′が生じたときの、事象Ｅの条件付き確率を以下の式で定義する。

$$P(E \mid E') = \frac{P(E \cap E')}{P(E')} \iff P(E \cap E') = P(E \mid E')P(E')$$

乗法公式：

$$P(E \cap E') = P(E \mid E')P(E')$$

事象の独立性:

２つの事象Ｅ，Ｅ′が独立とは、以下のことと等価である。

$$P(E \mid E') = P(E)$$

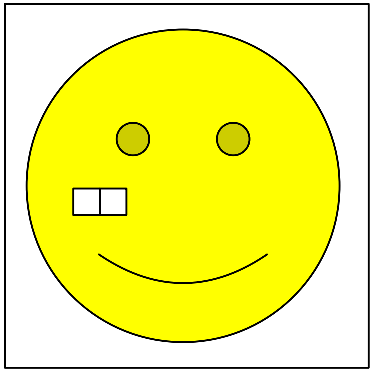
事象が独立ならば、以下の式が成り立つ。

$$P(E \cap E') = P(E \mid E')P(E') = P(E)P(E')$$

事象の独立性・従属性

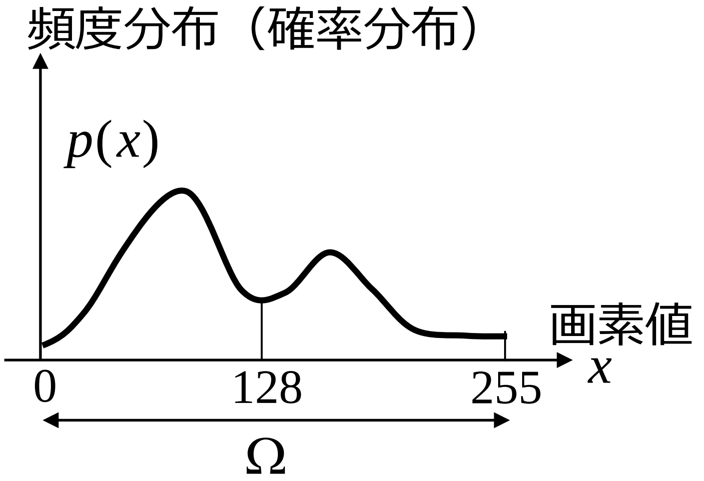
サイコロを２回振って、１回目に５、２回目に３が出る確率：

$$
\begin{aligned}
P(x_1=5,x_2=3) &= P(x_1=5)P(x_2=3\,|\,x_1=5)\\
&= P(x_1=5)P(x_2=3)\\
&= \frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36}
\end{aligned}
$$



ポートレイト画像中の２つの並んだ画素で、
左側が128,右側が127になる確率：

$$
\begin{aligned}
&P(x_1=128,x_2=127)\\
&\quad= P(x_1=128)P(x_2=127\,|\,x_1=128)\\
&\quad= P(x_1=128)P(x_2=127)?
\end{aligned}
$$



確率変数と確率分布

確率変数：
確率的にその値が生じるもの。あるいは試行によってはじめて値がわかるもの。
例）　サイコロを１回振って出る目：

$$P(X=i)=\frac{1}{6}, i\in\{1,2,3,4,5,6\}$$

一般に標本空間　$\Omega=\{\omega_1,\omega_2,...,\omega_n\}$　に属する要素のうち、
確率変数Xが値ω_kをとる事象を$X=\omega_k$で表し、その確率を以下のように表す。

$$P(X=\omega_k)=p_k, k=1,2,...,n \qquad （１）$$

式（１）を確率変数Ｘの従う確率分布とよぶ。確率の定義から

$$p_k\geq 0, k=1,2,...,n$$

および

$$p_1+p_2+\cdots+p_n=1$$

が成り立つ。

平均

確率変数Xの平均

大文字X：確率変数の総称、全体
小文字x: ある具体的な確率変数

有限個の実数値$\Omega = \{x_1, x_2, \ldots, x_n\}$をとる
確率変数Xの従う確率分布が

$$P(x_k) = p_k, k = 1,2,\ldots,n$$

で与えられるとき、確率分布Xの平均$E[X]$を

$$E[X] = \sum_{x\in\Omega} xP(x) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n$$

$$:= \mu$$ 文字の定義

によって表す。

関数f(X)の平均

$f(X)$の平均$E[f(X)]$は以下のように表される。

$$E[X] = \sum_{x\in\Omega} f(x)P(x) = f(x_1)p_1 + f(x_2)p_2 + \cdots + f(x_n)p_n$$

分散

関数$f(X)$を以下のように定義する。

$$f(x) := (x-\mu)^2$$

この関数の平均値は以下のように計算される。この量を分散という。

$$E[(X-\mu)^2] = \sum_{x\in\Omega}(x-\mu)^2 P(x)$$

$$= (x_1-\mu)^2 p_1 + (x_2-\mu)^2 p_2 + \cdots + (x_n-\mu)^2 p_n$$

$$:= \sigma^2$$ ⇦ 定義

分散は以下のようにも書ける。

$$E[(X-\mu)^2] = \sum_{x\in\Omega}(x-\mu)^2 P(x) = \sum_{x\in\Omega}(x^2 - 2\mu x + \mu^2)P(x)$$

$$= \sum_{x\in\Omega} x^2 P(x) - 2\mu \sum_{x\in\Omega} xP(x) + \mu^2 \sum_{x\in\Omega} P(x)$$

$$= \sum_{x\in\Omega} x^2 P(x) - 2\mu\cdot\mu + \mu^2\cdot 1$$

$$= \sum_{x\in\Omega} x^2 P(x) - \mu^2 = E[X^2] - \mu^2$$

離散系と連続系

離散系

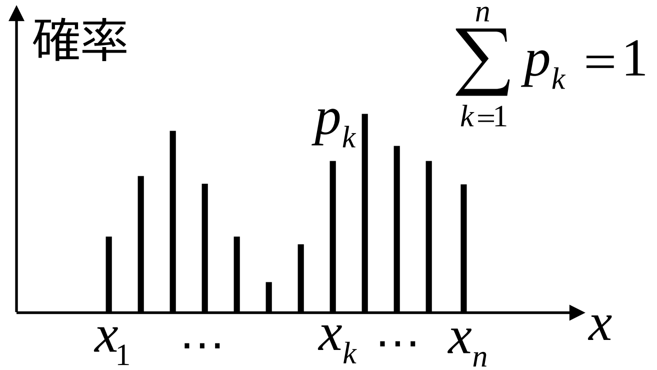
標本空間： $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$

確率分布： $P(x_k) = p_k, k = 1,2,\ldots,n$

確率変数の平均：

$$E[X] = \sum_{x \in \Omega} xP(x) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n$$

$$:= \mu$$



$\Omega = \{x_1, x_2, \ldots, x_n\}$

連続系

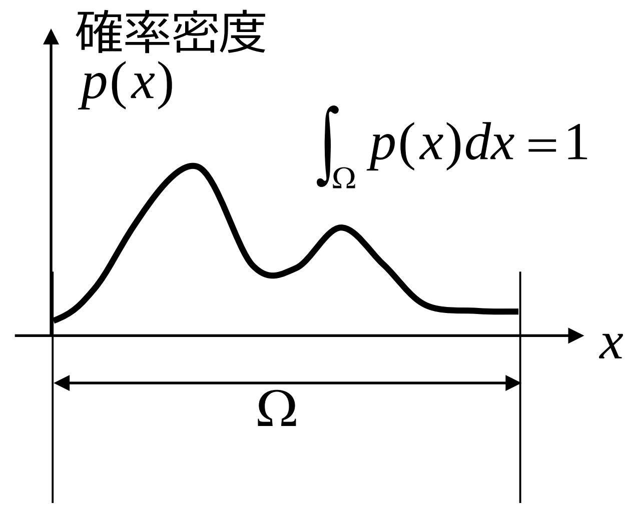
標本空間： $\Omega = \{\omega \,|\,$ある範囲の連続量$\}$

確率分布：$P(x) = p(x), x \in \Omega$

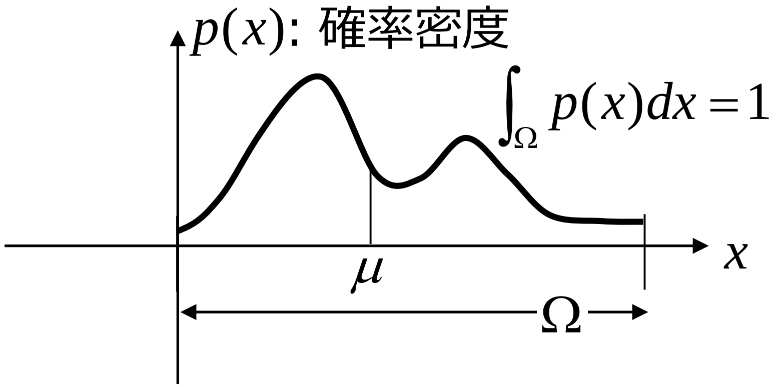
確率変数の平均：

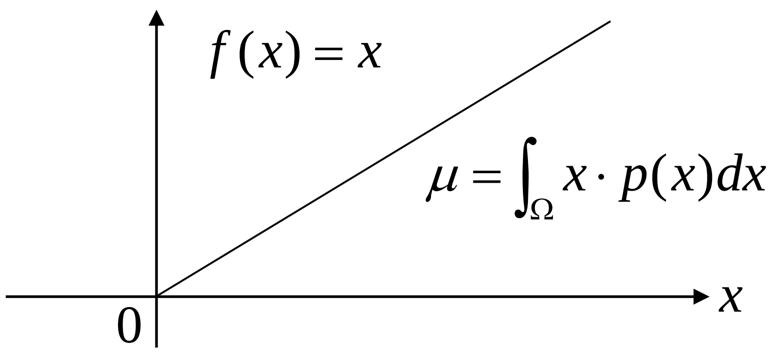
$$E[X] = \int_\Omega xP(x)dx$$

$$:= \mu$$

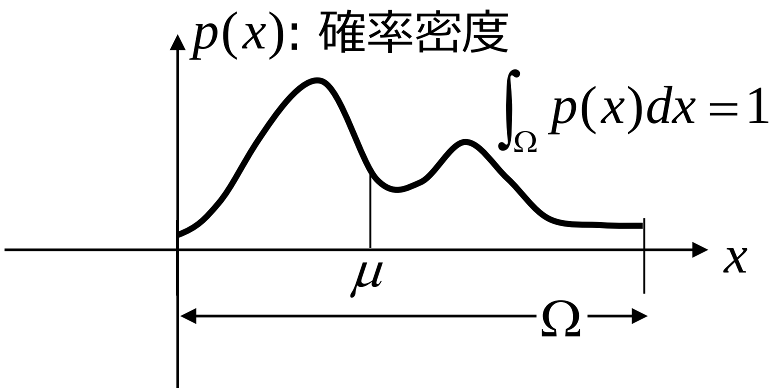


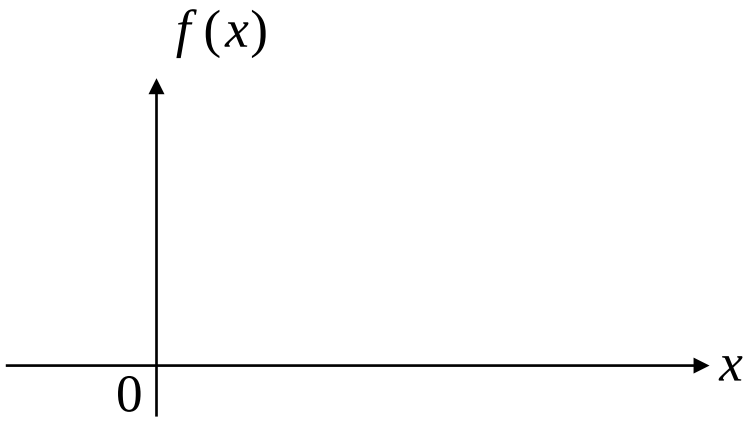
図による説明：平均


$p(x)$: 確率密度
$\int_{\Omega} p(x)dx = 1$
x
μ
Ω

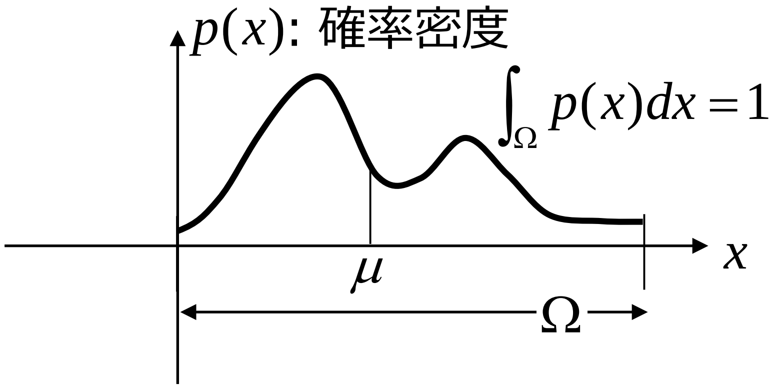

$f(x) = x$
$\mu = \int_{\Omega} x \cdot p(x)dx$
x
0

図による説明：分散





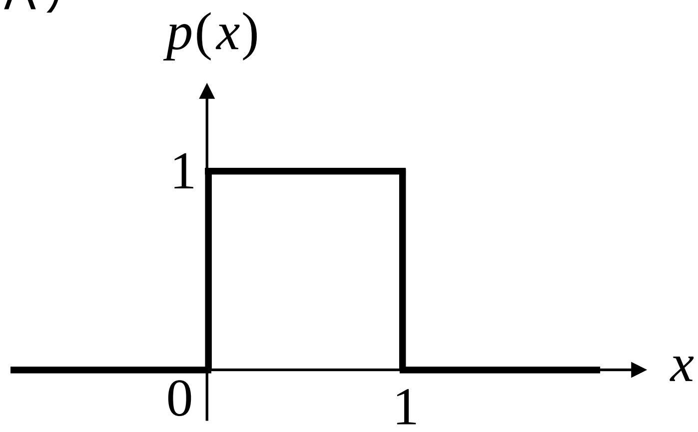
図による説明：エントロピー（平均情報量）



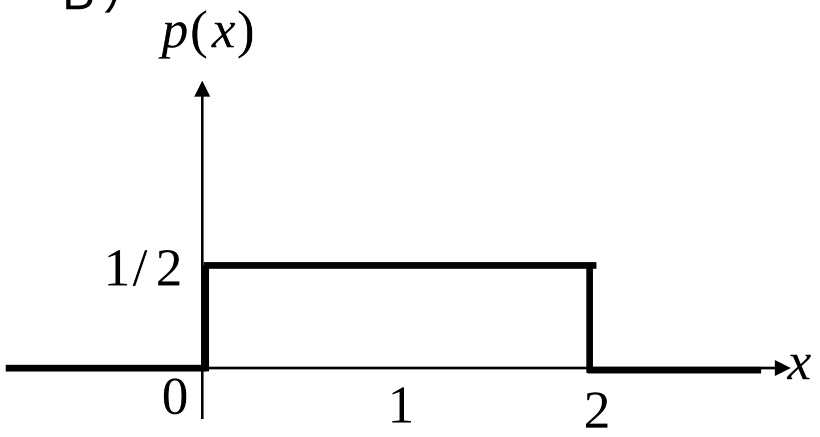
演習：分散の計算

次の２つ確率密度分布の場合について、確率変数xの平均を分散を求めなさい。

A）



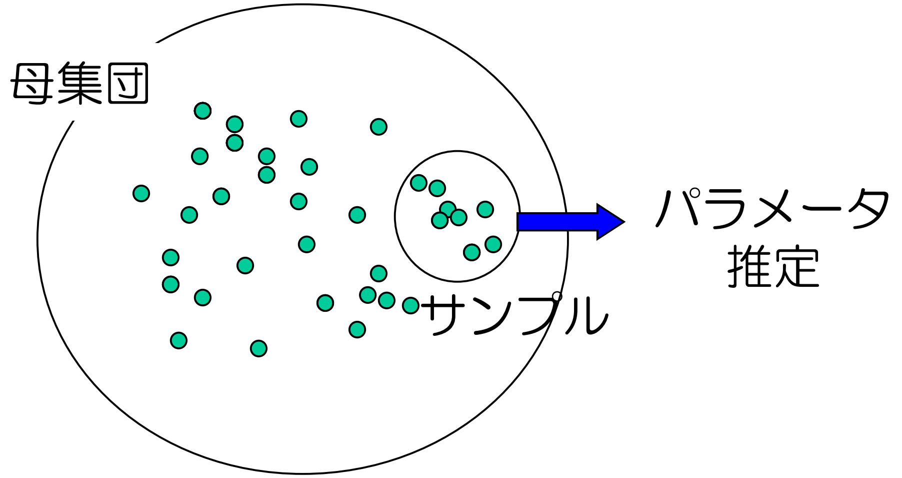
B）



補足資料

標本（サンプル）と母集団

データの解析、予測などの分野では、得られたサンプルから、その発生源である母集団の特性（平均、分散、…）を知りたい場合がほとんどである



例）
新しく開発された治療薬の有効性を評価したい。
身体の状態を表す数値を５０人の患者で計測し、
その数値から患者母集団全体の数値を推定したい。

標本から母集団の統計量を推定する

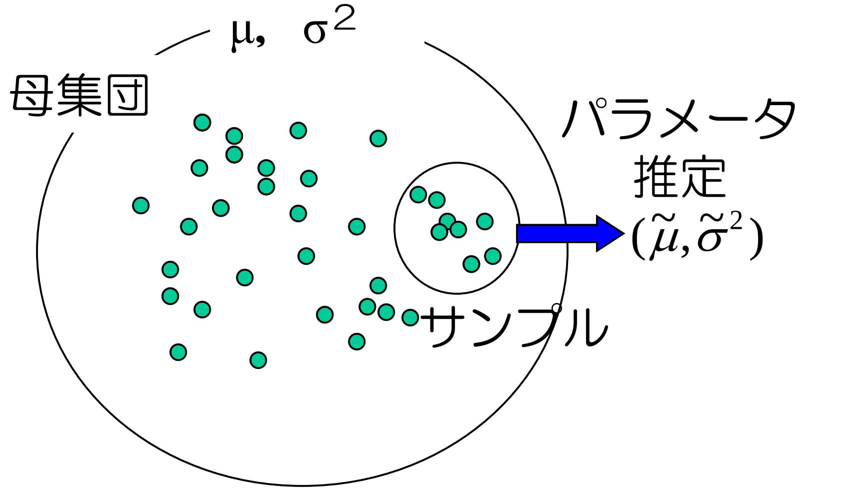
命題：
得られたサンプルから，
その発生母体である母集団の統計量を推定したい．

平均：１次の統計量

分散：２次の統計量

例）母集団が正規分布の場合

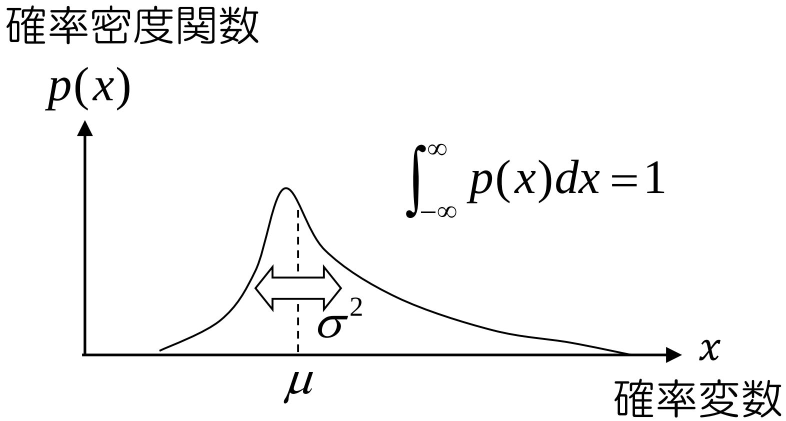
母集団を表すパラメータは平均μと分散σ^2
のふたつである．

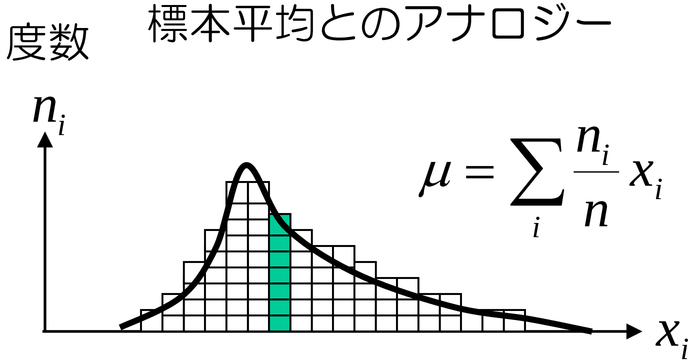


平均： $\mu = \int_{-\infty}^{\infty} \underbrace{x}_{1次} \cdot f(x)dx$

分散：$\sigma^2 = \int_{-\infty}^{\infty} \underbrace{(x-\mu)^2}_{2次} \cdot f(x)dx$

標本の平均、母集団の平均





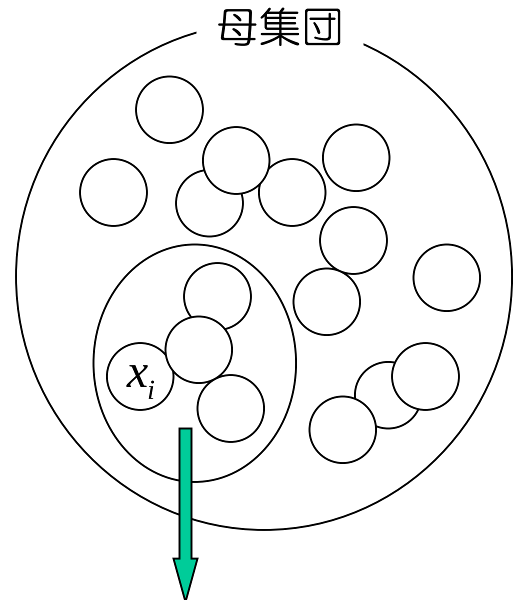
例）ある場所，ある日時での気温の確率．
x：気温
$p(x)$：気温xが起こる確率

中心極限定理 central limit theorem

分布がどのようなものであっても，平均値μ，分散σ^2を
もつ母集団からとられた大きさｎの標本の平均値の分布は，
ｎが大きくなるとき，正規分布$N(\mu, \sigma^2/n)$に近づく．
したがって，

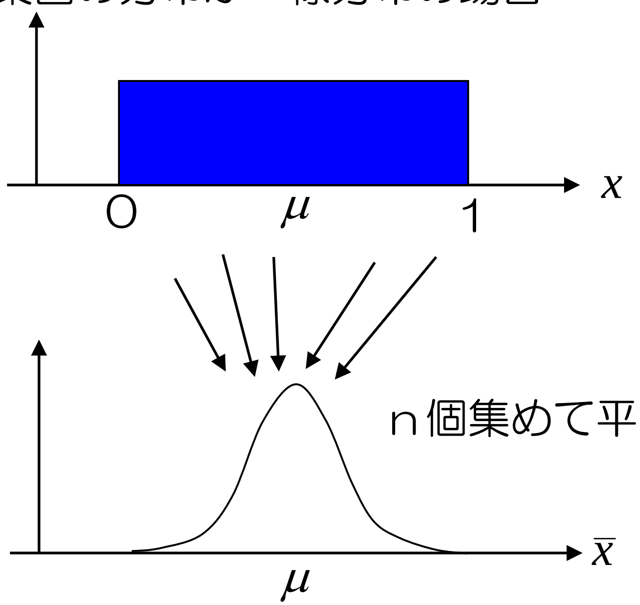
$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

の分布は，ｎが大となるとき，標準正規分布に近づく．

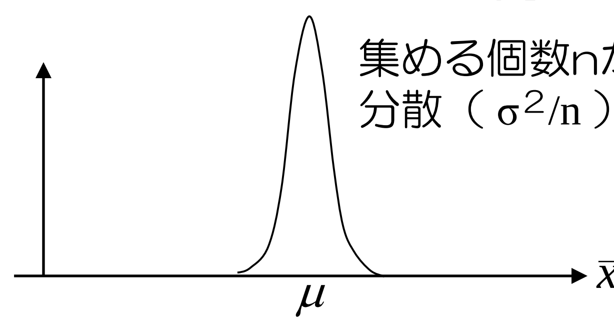


$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

例）母集団の分布が一様分布の場合



ｎ個集めて平均



集める個数nが多いほど
分散（σ^2/n）は小さい．

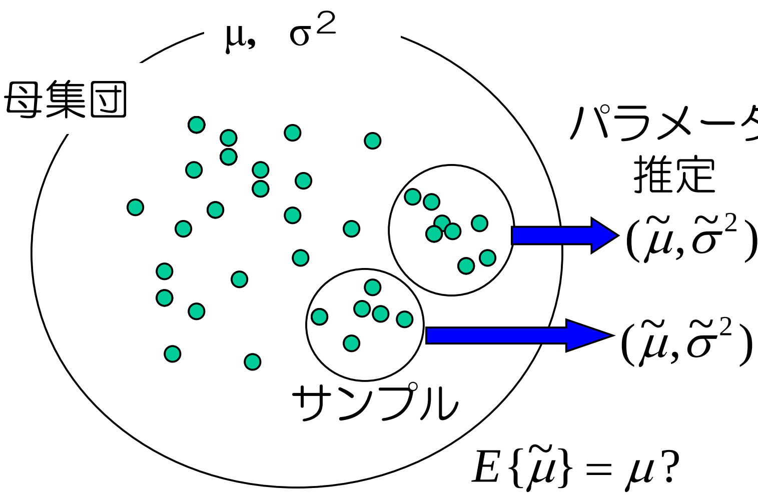
中心極限定理：多くの観測値を正規分布で近似する裏付けとなっている

不偏推定量 unbiased estimator 平均の不偏推定量

不偏推定量とは，サンプルから求めた
母集団パラメータの期待値が，真の
母集団パラメータに一致するものをいう．

例）母集団が正規分布の場合

母集団を表すパラメータは平均μと分散σ^2
のふたつである．



$$E\{\tilde{\mu}\} = \mu\,?$$

$$E\{\tilde{\sigma}^2\} = \sigma^2\,?$$

仮定：ひとつひとつのサンプルは独立であり、また同じ確率分布にしたがって発生するとする independent and identically distributed random variables (i.i.d.モデル)

母集団平均の推定をサンプル平均で行った場合，

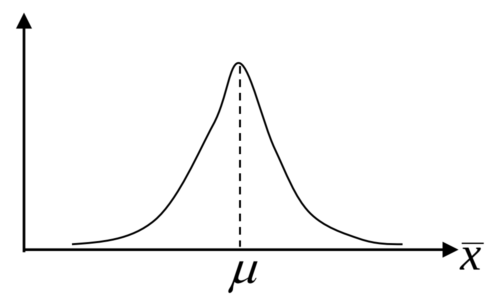
$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i$$

サンプル平均の期待値は

$$E\{\bar{x}\} = E\left\{\frac{1}{m}\sum_{i=1}^{m} x_i\right\} = \frac{1}{m}\sum_{i=1}^{m} E\{x_i\}$$

$$= \frac{1}{m}\sum_{i=1}^{m} \mu = \frac{m}{m}\mu = \mu$$

となり，母集団平均に一致する．
よって，サンプル平均は，母集団
平均に対する不偏推定量といえる．



分散の不偏推定量

標本分散の期待値を計算してみる　$s^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \bar{x})^2$

$$E\left\{\frac{1}{m}\sum_{i=1}^{m}(x_i - \bar{x})^2\right\} = E\left\{\frac{1}{m}\sum_{i=1}^{m}[(x_i - \mu) - (\bar{x} - \mu)]^2\right\}$$

$$= \frac{1}{m}E\left\{\sum_{i=1}^{m}(x_i - \mu)^2\right\} - \frac{2}{m}E\left\{\sum_{i=1}^{m}(x_i - \mu)(\bar{x} - \mu)\right\} + \frac{1}{m}E\left\{\sum_{i=1}^{m}(\bar{x} - \mu)^2\right\}$$

$$= \boxed{\frac{1}{m}E\left\{\sum_{i=1}^{n}(x_i - \mu)^2\right\} - \frac{2}{m}E\left\{\sum_{i=1}^{m}(x_i - \mu)(\bar{x} - \mu)\right\} + \frac{1}{m}E\left\{m(\bar{x} - \mu)^2\right\}}$$

上式右辺の第１項は

$$\frac{1}{m}E\left\{\sum_{i=1}^{m}(x_i - \mu)^2\right\} = \frac{1}{m}\sum_{i=1}^{m}E\{(x_i - \mu)^2\} = \frac{1}{m}\sum_{i=1}^{m}\sigma^2 = \sigma^2$$

分散の不偏推定量

第３項は

$$E\{(\bar{x}-\mu)^2\}=E\left\{\left(\frac{1}{m}\sum_{i=1}^{m}x_i-\mu\right)^2\right\}=E\left\{\left(\frac{1}{m}\sum_{i=1}^{m}x_i-\frac{1}{m}\sum_{i=1}^{m}\mu\right)^2\right\}$$

$$=E\left\{\left(\frac{1}{m}\sum_{i=1}^{m}(x_i-\mu)\right)^2\right\}$$

$$=E\left\{\frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m}(x_i-\mu)(x_j-\mu)\right\}$$

$$=E\left\{\frac{1}{m^2}\sum_{i=1}^{m}(x_i-\mu)^2\right\}$$

$$=\frac{1}{m^2}\sum_{i=1}^{m}(E\{x_i\}-\mu)^2$$

$$=\frac{1}{m^2}m\sigma^2=\frac{\sigma^2}{m}$$

分散の不偏推定量

第２項は：演習

分散の不偏推定量（つづき）直感的解釈

なぜ分散の推定を，（mで割らずに）

$$\hat{\sigma}^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2$$

で与えるか？ ⇒直感的解釈

仮に母集団の平均μが既知であれば，
m個のデータからの分散の推定は

$$\bar{\sigma}^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)^2$$

で与えればよい．これに対し，母集団平均μ
が未知のために，かわりにサンプル平均を
用いた場合の分散をｓ^{2}とすると，

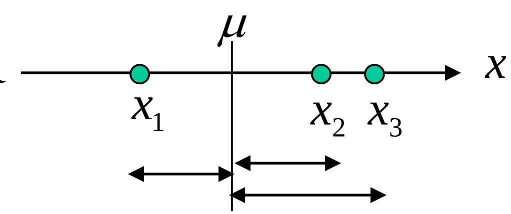
$$s^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \bar{x})^2$$

この場合，かならず

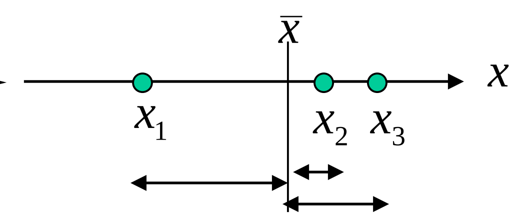
$$s^2 \le \bar{\sigma}^2$$

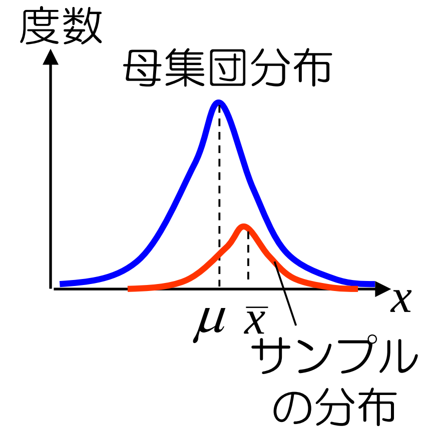
が成り立つ．すなわち，ｓ^{2}は真の
母集団分散を過小に推定する傾向がある．
そこで，mで割らずにm－１で割ることで
この過小推定を防ぐ．

真の母集団平均



サンプルから
求めた平均





サンプルから母集団の平均を推定する

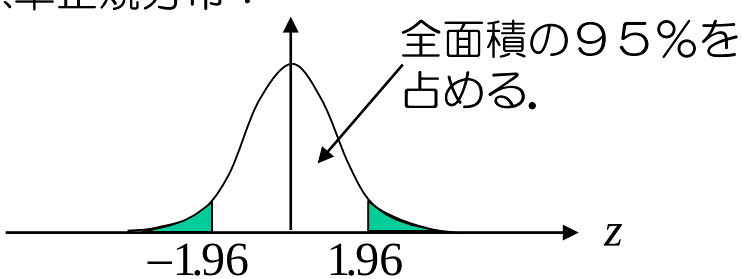
母集団が正規分布に従うとする.
もし母集団正規分布の平均と分散が既知なら
ｎ個のサンプルを集めてきて得た平均$\bar{x}$は
ｎの値によらず，正規分布$N(\mu, \sigma^2/n)$をする.

標準化（平均を引き，標準偏差で割る）を
行えば，その値は標準正規分布に従う.

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

いま，母集団の分散σ^2のみが既知としたとき，
標本から推定される母集団平均μの区間を
考える.

標準正規分布：



標準正規分布は－１．９６から１．９６の
間をとる確率が９５％である.

$$P\left\{-1.96 < \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} < 1.96\right\} = 0.95$$

カッコの中を書き直せば，

$$P\left\{-\bar{x} - 0.96\frac{\sigma}{\sqrt{n}} < -\mu < -\bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right\} = 0.95$$

$$P\left\{\bar{x} - 0.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right\} = 0.95$$

これより，未知の母集団平均μが

$$\left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$$

という範囲に９５％の確率で存在する
ことがわかる.

正規分布（ガウス分布）

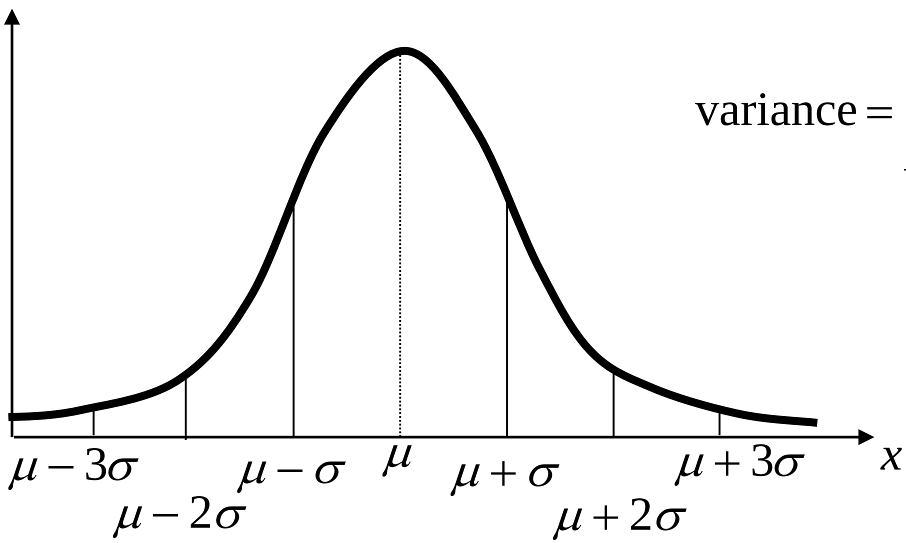
確率密度関数

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

この分布の平均と分散は,

$$mean=\int_{-\infty}^{\infty}xf(x)dx=\int_{-\infty}^{\infty}x\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx=\mu$$

$$\mathrm{variance}=\int_{-\infty}^{\infty}(x-\mu)^2\cdot f(x)dx=\sigma^2$$



正規分布は平均μと分散σ２によって完全に記述される.

⇩

$N(\mu,\sigma^2)$ と表記する

確率変数の範囲と確率（よく用いられる値）

$\mu-\sigma\le x\le\mu+\sigma$ ----- 68.27%

$\mu-2\sigma\le x\le\mu+2\sigma$ ----- 95.45%

$\mu-3\sigma\le x\le\mu+3\sigma$ ----- 99.73%

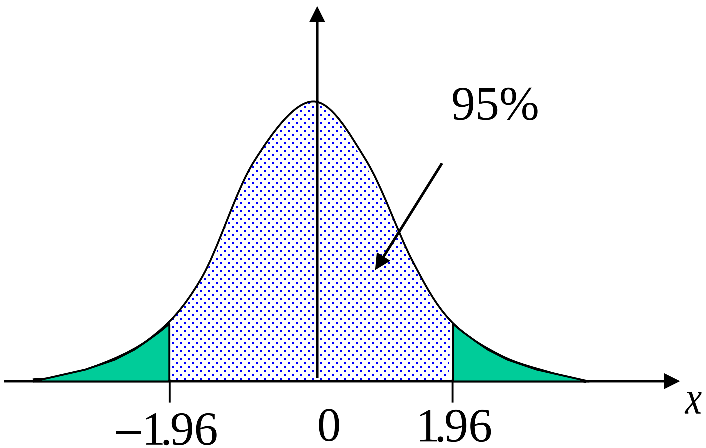
$\mu-1.96\sigma\le x\le\mu+1.96\sigma$ ----- 95%

特に，平均０，分散１の正規分布N(0,1)を標準正規分布と呼ぶ.

正規分布（ガウス分布）つづき

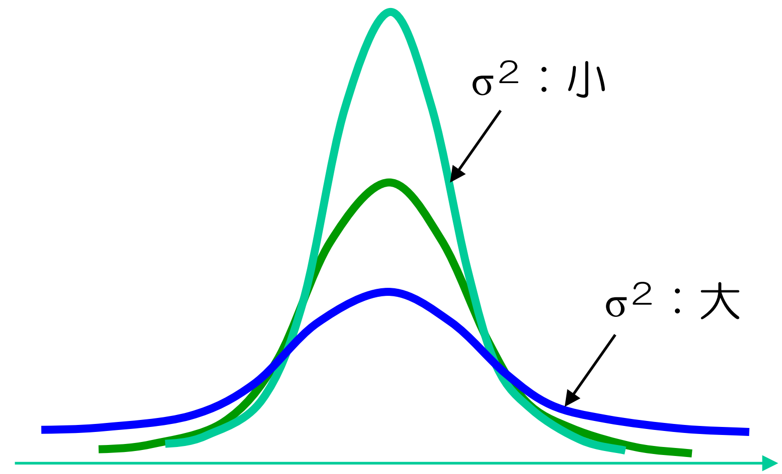
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

標準正規分布 $N(0,1)$

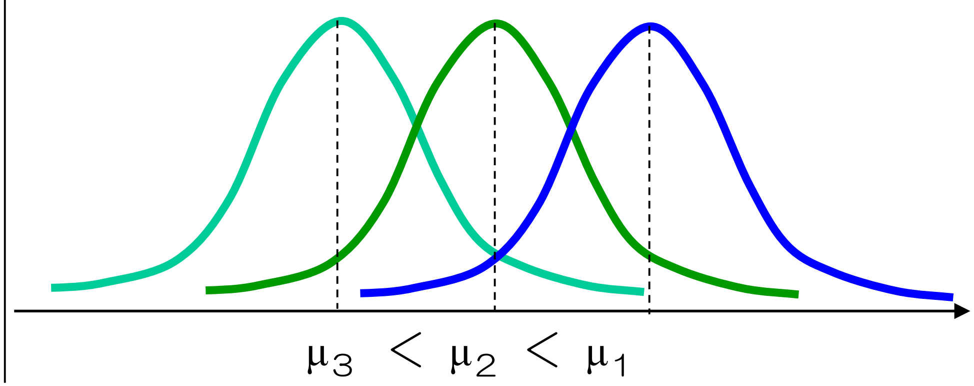


９５％の確率で存在する範囲が統計ではしばしば使われる。
標準正規分布では－１．９６から１．９６の範囲となる。

平均が同じで分散が異なる正規分布



分散が同じで平均が異なる正規分布



二項分布binomial distribution

例）３回サイコロを投げて，ｘ回，１の目が出る確率を考える.

	０回	１回	２回	３回
$P(x)$	$\left(\frac{5}{6}\right)^3$	$3\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^2$	$3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)$	$\left(\frac{1}{6}\right)^3$
	$\boxed{\neq 1}\boxed{\neq 1}\boxed{\neq 1}$	$\boxed{1}\boxed{\neq 1}\boxed{\neq 1}$	$\boxed{1}\boxed{1}\boxed{\neq 1}$	$\boxed{1}\boxed{1}\boxed{1}$

$\Rightarrow$ $P(x)={}_3C_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{3-x}$

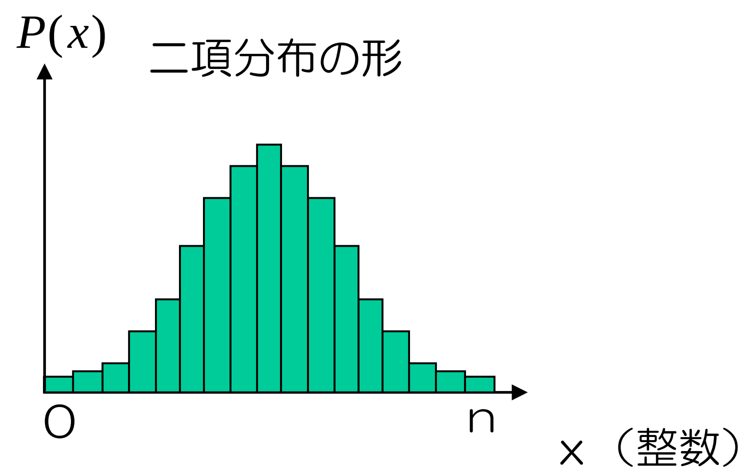
一般に，確率ｐをもつ事象が，
ｎ回の観察でｘ回起こる確率P(x)は

$$P(x)={}_nC_x p^x(1-p)^{n-x}=\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

この式で表される確率分布を二項分布と呼ぶ.

平均： $\mu = np$

分散： $\sigma^2 = np(1-p)$



ｎが大きくなると，二項分布は
正規分布に近づく

ポアソン分布Poisson distribution

二項分布において，実験回数ｎが十分大きい場合，
二項分布はポアソン分布で近似できる．

$$P(x) = {}_nC_x p^x (1-p)^{n-x}$$

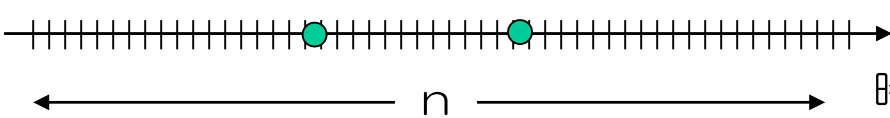
⇩ 近似

$$P(x) = \frac{\mu^x e^{-\mu}}{x!} \qquad \text{ただし} \quad \mu = np$$

平均μが大きければ，ポアソン分布は
正規分布に近似できる．

例）千葉市の１日あたりの交通事故件数の分布

１日を十分細かくきざんで考える（例えば１分単位）．
すると，このきざみのなかでは，事故が起こるか起こらないかの，どちらかの事象のみ起こるとみなせる．
１つのきざみ内で事故が起こる確率をｐとすれば，
１日にｘ件事故が起こる確率は，二項分布で表せる．



１日平均５回，事故が起こるとする．

１）二項分布で考えると，

１分あたりに事故が起こる確率は

$$p = 5/(24 \times 60)$$

ある１日に，ｘ回起こる確率は，

$$P(x) = {}_{24\times 60}C_x p^x (1-p)^{24\times 60 - x}$$

２）ポアソン分布で考えると

$$P(x) = \frac{5^x e^{-5}}{x!}$$

事故数	二項分布	ポアソン分布
0	0.00668	0.00674
1	0.03351	0.03369
2	0.08402	0.08422
3	0.14032	0.14037
4	0.17565	0.17547
5	0.17577	0.17547
6	0.14648	0.14622
7	0.10455	0.10444
8	0.06526	0.06528
9	0.03618	0.03627
10	0.01804	0.01813

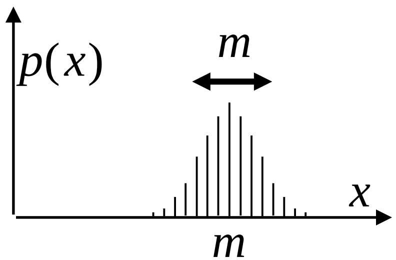
５回

ポアソン分布の性質とフォトンノイズの例

ポアソン分布は，平均と分散が等しい．

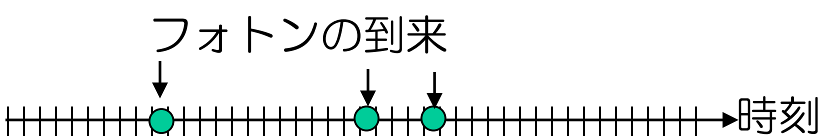
$$P(x) = \frac{m^x e^{-m}}{x!}$$

において 平均＝分散＝m

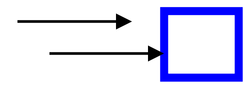
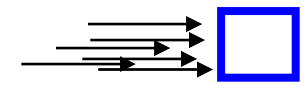


標準偏差は $\sigma = \sqrt{m}$

例）明るい条件と暗い条件で，単位時間あたりにCCDの画素に到達するフォトン数を考える．



CCDの画素に到達するフォトン数はポアソン分布に従う．

[暗い]	[明るい]
 CCD画素	 CCD画素
平均をm=100とする	平均をm=10000とする

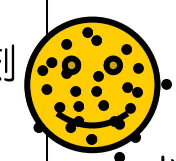
標準偏差は

$\sigma = \sqrt{100} = 10$	$\sigma = \sqrt{10000} = 100$

フォトン数ｘのちらばりを
±２σの範囲で考えると

$80 < x < 120$	$9800 < x < 10200$

カメラのゲインコントロールによって
明るさを合わせられることを考えて，
それぞれの平均が１００になるように
正規化すると

 $80 < x < 120$	$98 < x < 102$ 

以上より，暗い状態ではノイズが増える
ことがわかる（フォトンノイズという）