

九州大学 工学部地球環境工学科
船舶海洋システム工学コース

海事統計学 第5回 (担当:木村)

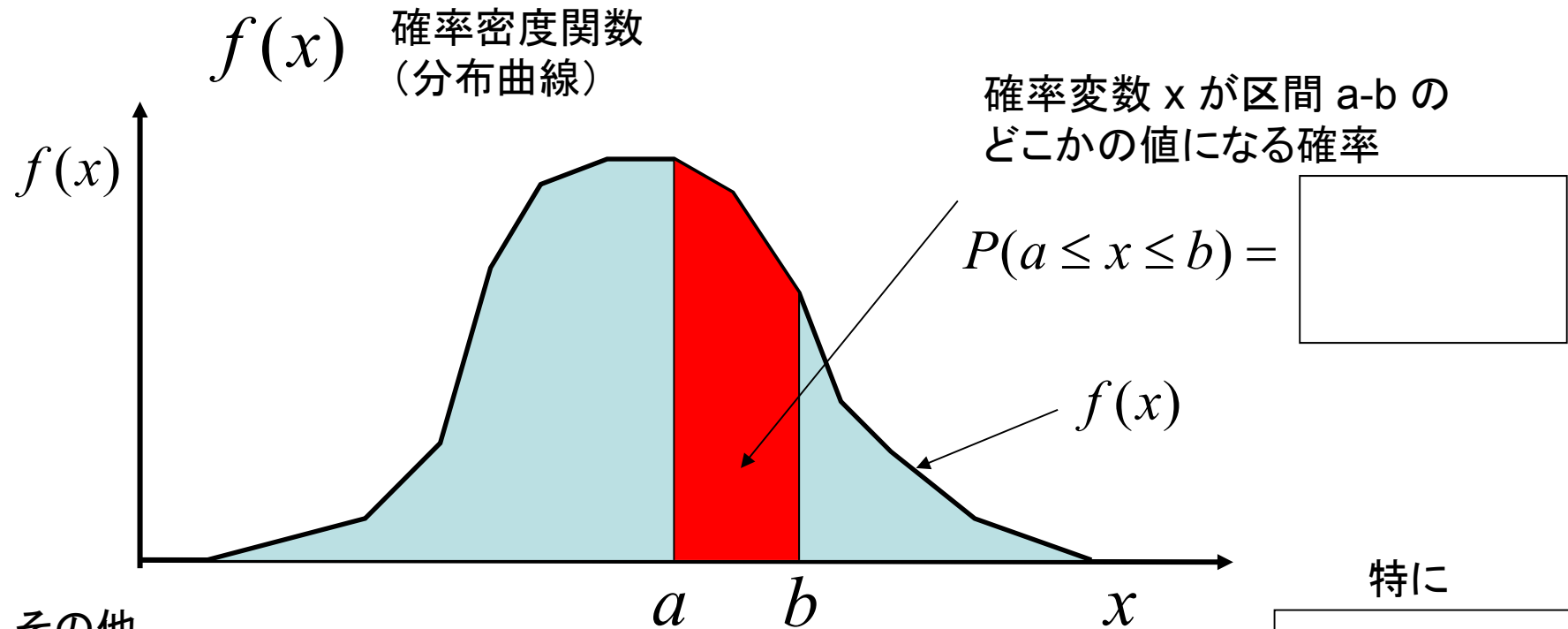
確率論の基礎3
(確率変数が連続値をとる場合)

授業の資料等は

<http://sysplan.nams.kyushu-u.ac.jp/gen/index.html>

確率変数のとりうる値が連続値の場合

- ・連続 (continuous) 確率変数
- ・確率変数 x がある特定の値をとる「確率」はゼロ
- 「確率」は、確率変数のとりうる値の任意の範囲に対して割当てられる
- ・確率変数 x のそれぞれの値に対して**確率密度 (probability density)** が対応



その他、

$$P(-\infty \leq x \leq \infty) = \int_{-\infty}^{\infty} f(x) dx = 1 \quad (\text{累積}) \text{分布関数は1に収束}$$

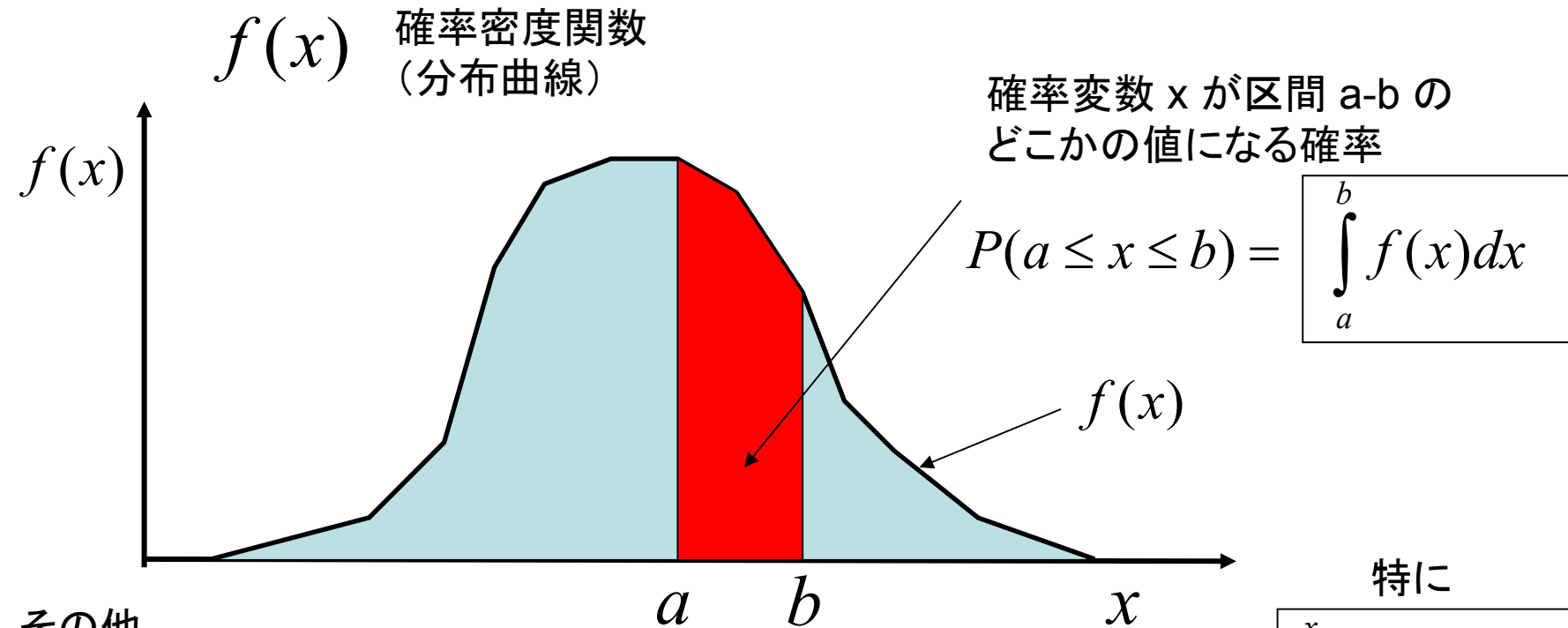
$f(x) \geq 0$ 確率密度関数は非負 だが $f(x) \leq 1$ とは限らない

特に

(累積) 分布関数

確率変数のとりうる値が連続値の場合

- ・連続 (continuous) 確率変数
- ・確率変数 x がある特定の値をとる「確率」はゼロ
- 「確率」は、確率変数のとりうる値の任意の範囲に対して割当てられる
- ・確率変数 x のそれぞれの値に対して**確率密度 (probability density)** が対応



その他、
$$P(-\infty \leq x \leq \infty) = \int_{-\infty}^{\infty} f(x) dx = 1$$
 (累積) 分布関数は 1 に収束

$f(x) \geq 0$ 確率密度関数は非負 だが $f(x) \leq 1$ とは限らない

特に

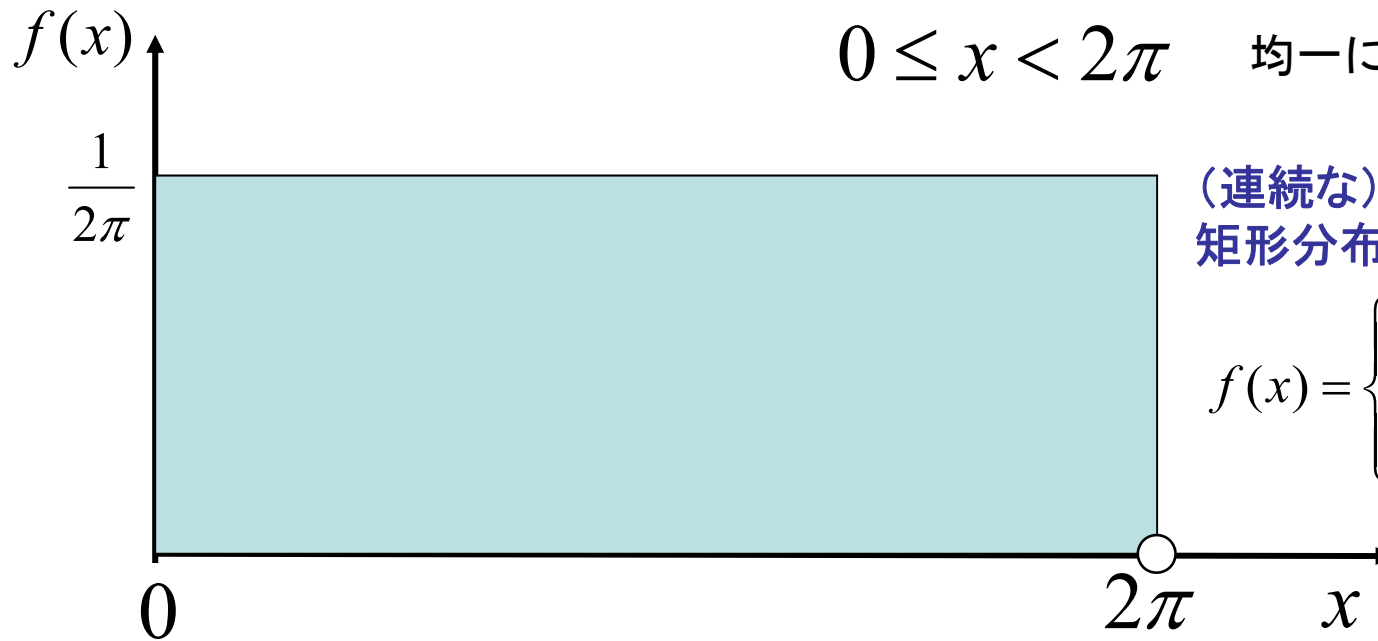
$$\int_{-\infty}^x f(y) dy = F(x)$$

(累積) 分布関数

確率変数が連続値の例

1本の針を投げたとき、ある基準線となす角度 x

$$0 \leq x < 2\pi \quad \text{均一にどの角度にもなりうる}$$



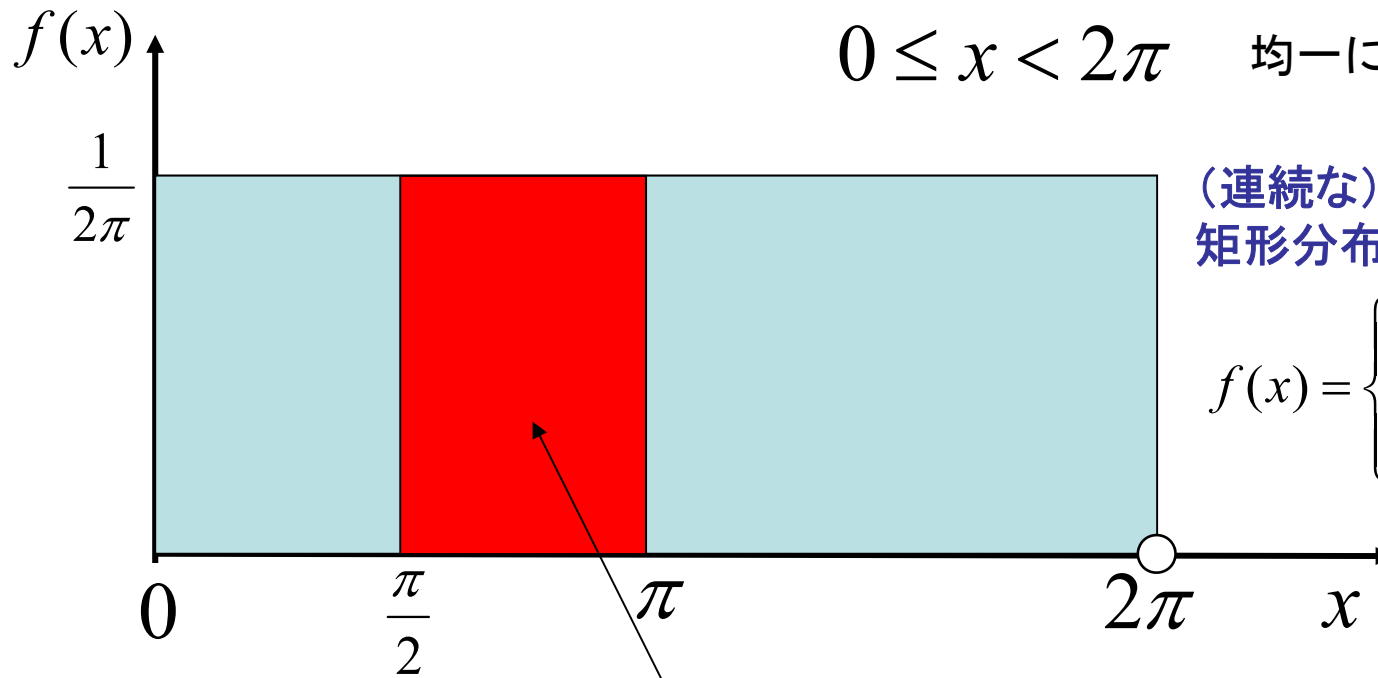
(連続な) 一様分布 矩形分布 コンピュータで簡単に生成可能

$$f(x) = \begin{cases} \frac{1}{2\pi} & \text{while } 0 \leq x < 2\pi \\ 0 & \text{otherwise} \end{cases}$$

【問1】 針の角度が $\frac{\pi}{2} \leq x \leq \pi$ となる確率は？

確率変数が連続値の例

1本の針を投げたとき、ある基準線となす角度 x
 $0 \leq x < 2\pi$ 均一にどの角度にもなりうる



(連続な) 一様分布 矩形分布 コンピュータで簡単に生成可能

$$f(x) = \begin{cases} \frac{1}{2\pi} & \text{while } 0 \leq x < 2\pi \\ 0 & \text{otherwise} \end{cases}$$

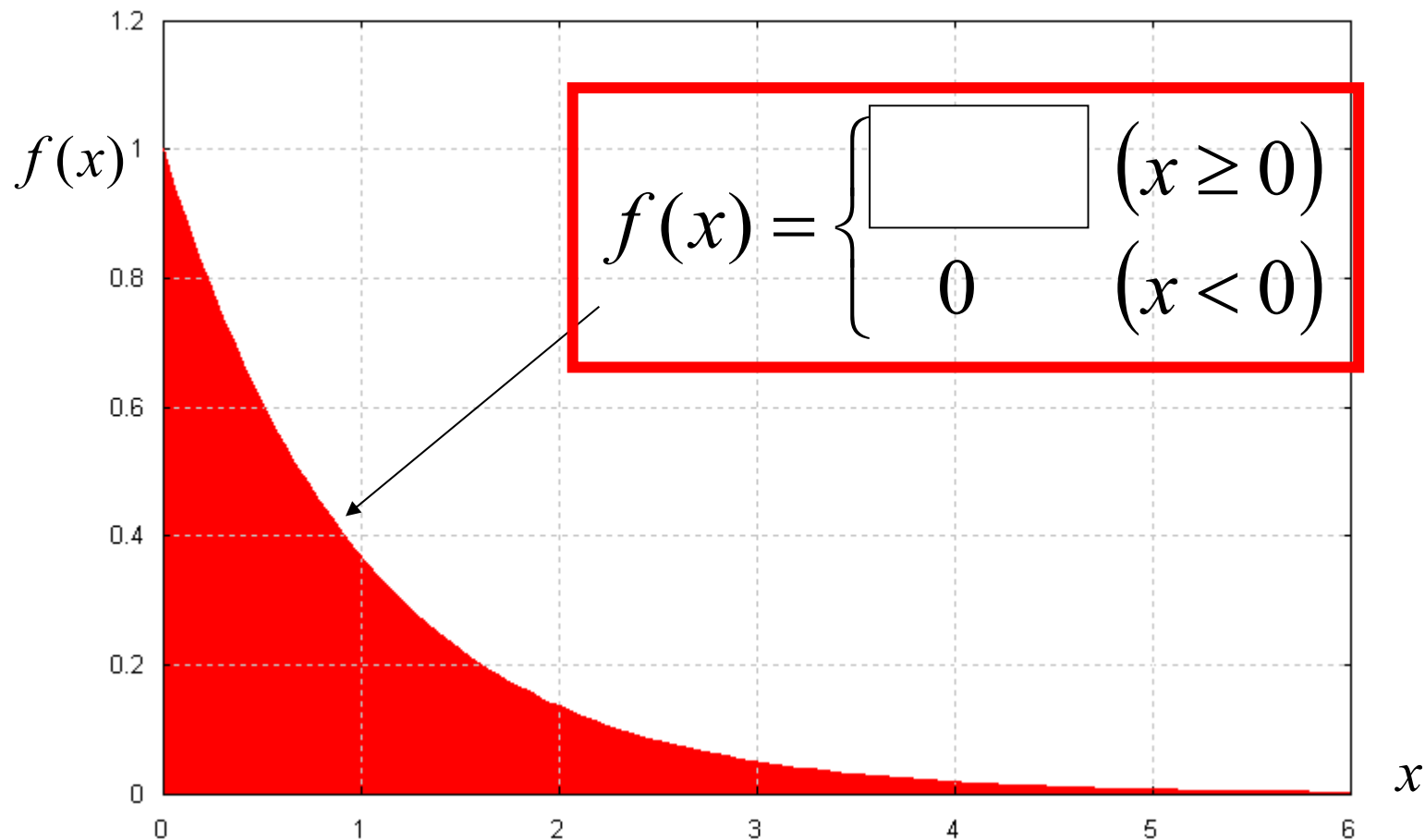
【問1】 針の角度が $\frac{\pi}{2} \leq x \leq \pi$ となる確率は？

$$P\left(\frac{\pi}{2} \leq x \leq \pi\right) = \int_{\frac{\pi}{2}}^{\pi} f(x) dx = \frac{1}{4}$$

代表的な連続形分布関数：指数分布

exponential distribution

機械が偶発的に故障するまでの時間の分布や
客が訪れる時間間隔などのモデルによく用いられる



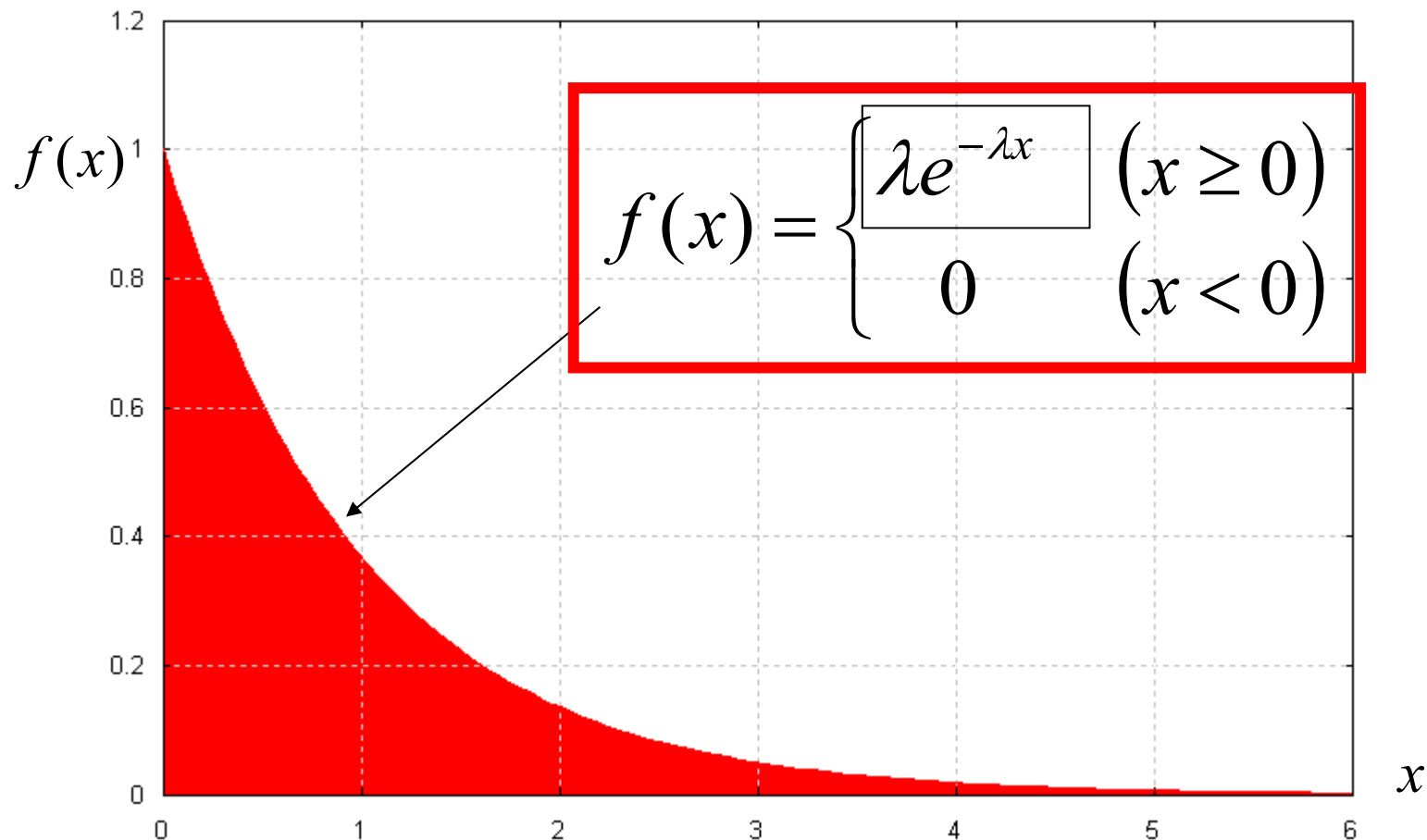
代表的な連続形分布関数：指数分布

exponential distribution

λ : 平均到着(故障)率(人/時)

機械が偶発的に故障するまでの時間の分布や
客が訪れる時間間隔などのモデルによく用いられる

$1/\lambda$: 平均到着(故障)時間間隔



【復習】 確率変数のとりうる値が離散かつ無限個の場合

ポアソン分布 (Poisson distribution)

発生時間間隔は
指数分布

単位時間あたり平均で λ 回発生する事象が、
単位時間に x 回(ゼロを含む)発生する確率

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

ポアソン分布の現れる例:

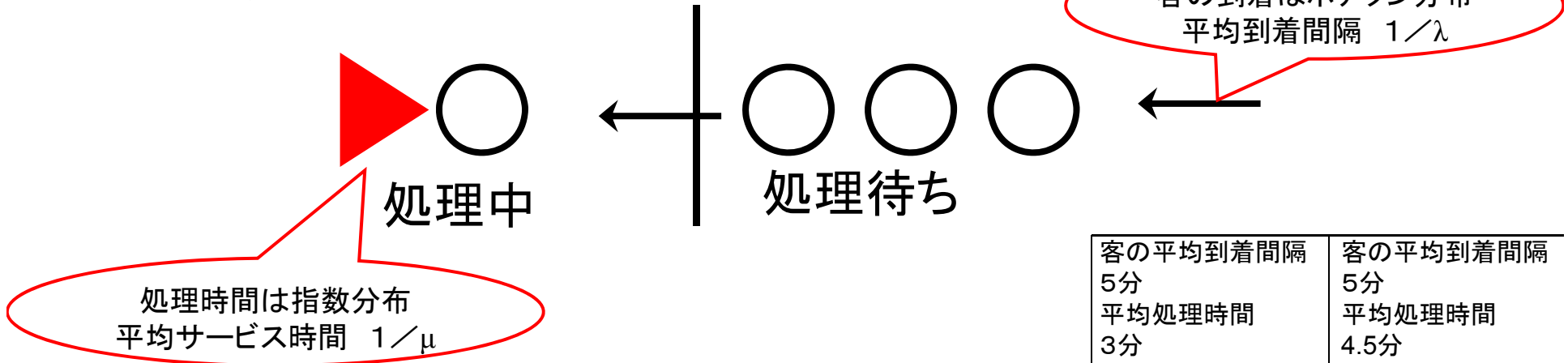
備考: $x = 0$ のとき $x! = 1$

- 1分間に放射性物質から放射される粒子が平均2個観測されるとき、1分間に1個も観測されない確率は？
- ある地域において、年間の交通事故件数が平均730件のとき、1日に発生する事故の件数が0件である確率は？
- ある機械で部品を 10000個作ると、平均 5個の不良品ができるとき、部品を 1000個作ったときに不良品が0個である確率は？ 不良品が1個出る確率は？

指数分布 と ポアソン分布 は同一現象の異なる表現

λ : 平均到着(故障)率(人/時) $\Leftrightarrow \frac{1}{\lambda}$: 平均到着(故障)時間間隔

【参考】指数分布・ポアソン分布の応用：待ち行列



客の平均到着間隔 5分 平均処理時間 3分	客の平均到着間隔 5分 平均処理時間 4.5分

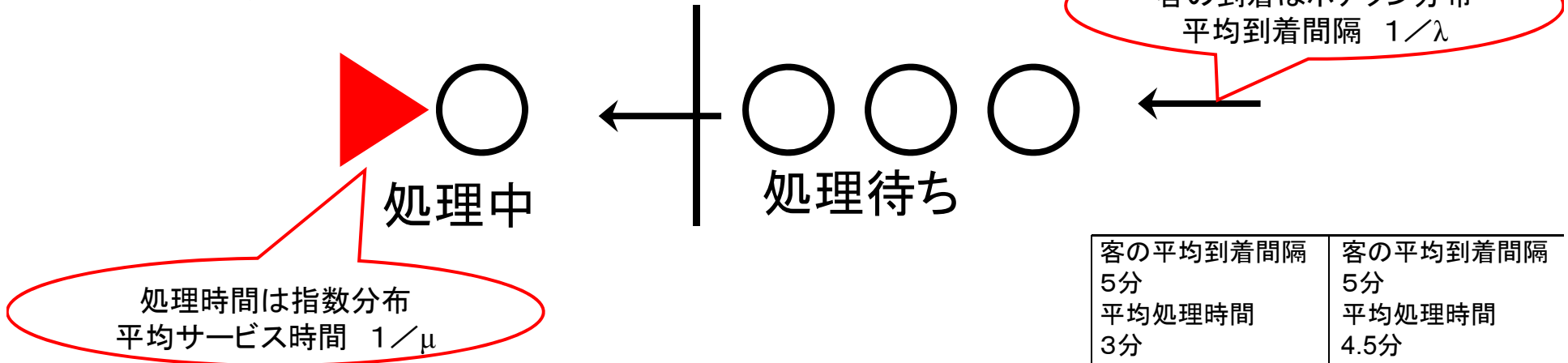
サービス中を含む客の数がゼロである確率
(待たずにサービスが受けられる確率) $1 - \frac{\lambda}{\mu}$

サービス中を含む客の平均人数

サービスを待っている客の平均人数

到着してからサービスを受けて去るまでの平均時間

【参考】指数分布・ポアソン分布の応用：待ち行列



客の平均到着間隔 5分 平均処理時間 3分	客の平均到着間隔 5分 平均処理時間 4.5分
1.5人	9人
7.5分	45分

サービス中を含む客の数がゼロである確率
(待たずにサービスが受けられる確率) $1 - \frac{\lambda}{\mu}$

サービス中を含む客の平均人数 $\frac{\frac{\lambda}{\mu}}{1 - \frac{\lambda}{\mu}}$

サービスを待っている客の平均人数 $\frac{\left(\frac{\lambda}{\mu}\right)^2}{1 - \frac{\lambda}{\mu}}$

到着してからサービスを受けて去るまでの平均時間 $\frac{1}{\mu - \lambda}$

待ち行列理論できちんと解析せずに設備投資をケチると...



トイレ行列・ATMの行列など

顧客へ長い待ち時間を強いる
顧客離れ
企業イメージの低下

→ ビジネスチャンスを逃す



ストックヤード：クレーンやコンベア能力によっては、貨物や資材で溢れ返ることに

期待値 (expected value)

確率分布関数や確率密度関数から計算される確率変数 x の(理論的な)平均値

離散分布の場合

$$E\{x\} = \sum_{x \in X} x P(x)$$

連続分布の場合

$$E\{x\} = \boxed{\phantom{E\{x\} = \int x f(x) dx}}$$

確率分布関数
(確率密度関数)
の重心

σ は標準偏差
と呼ばれる

分散 (variance)

確率変数 x についての期待値からの偏差の2乗の期待値

$$Var\{x\} = \sigma^2$$

離散分布の場合

$$Var\{x\} = E\{(x - E\{x\})^2\} = \sum_{x \in X} (x - E\{x\})^2 P(x)$$

連続分布の場合

$$Var\{x\} = E\{(x - E\{x\})^2\} = \boxed{\phantom{Var\{x\} = \int (x - E\{x\})^2 f(x) dx}}$$

期待値 (expected value)

確率分布関数や確率密度関数から計算される確率変数 x の(理論的な)平均値

離散分布の場合

$$E\{x\} = \sum_{x \in X} x P(x)$$

連続分布の場合

$$E\{x\} = \int_{-\infty}^{\infty} x f(x) dx$$

確率分布関数
(確率密度関数)
の重心

σ は標準偏差
と呼ばれる

分散 (variance)

確率変数 x についての期待値からの偏差の2乗の期待値

$$Var\{x\} = \sigma^2$$

離散分布の場合

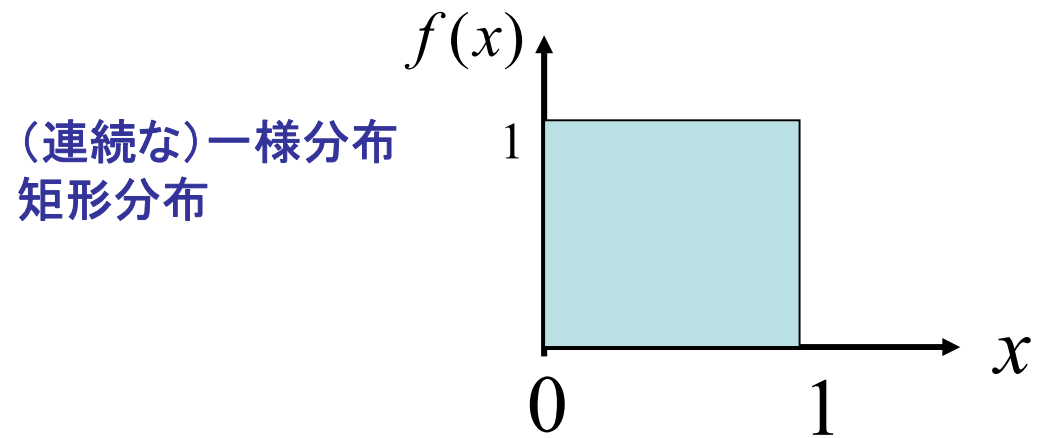
$$Var\{x\} = E\{(x - E\{x\})^2\} = \sum_{x \in X} (x - E\{x\})^2 P(x)$$

連続分布の場合

$$Var\{x\} = E\{(x - E\{x\})^2\} = \int_{-\infty}^{\infty} (x - E\{x\})^2 f(x) dx$$

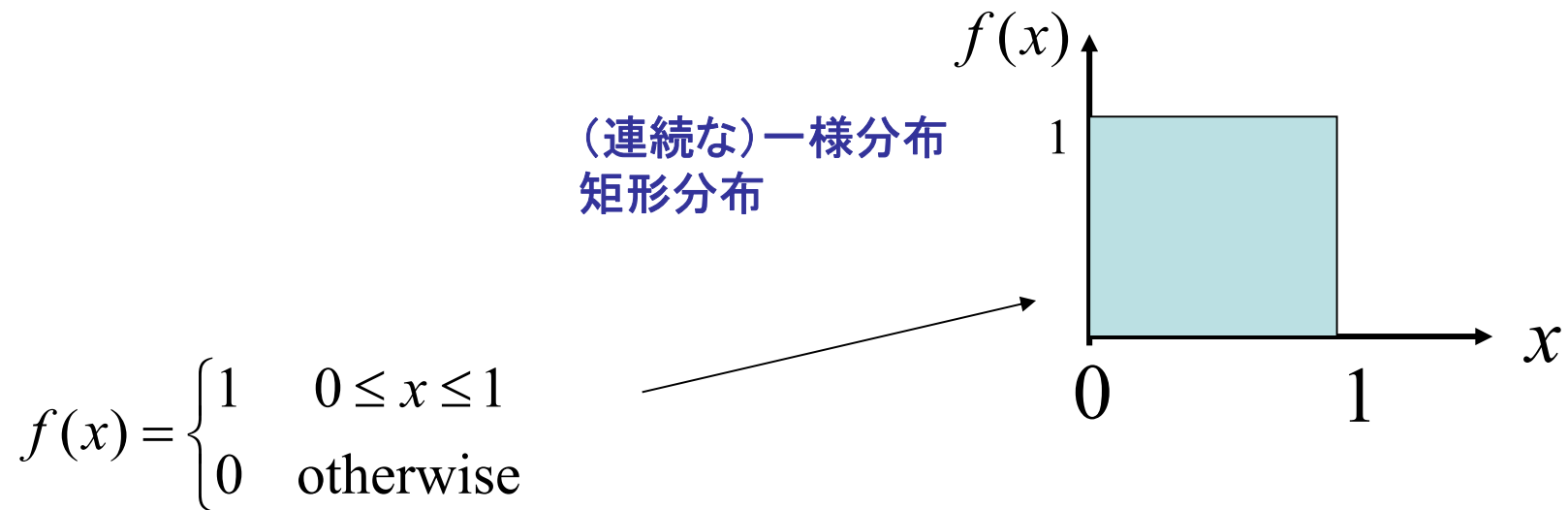
練習問題

【1】 $0 \leq x \leq 1$ の連続値をとる一様分布の期待値と分散を計算せよ。



練習問題

【1】 $0 \leq x \leq 1$ の連続値をとる一様分布の期待値と分散を計算せよ。



$$\mu = \int_0^1 xf(x)dx = \int_0^1 xdx = \left[\frac{x^2}{2} \right]_0^1 = \frac{1}{2}$$

$$\sigma^2 = \int_0^1 (x - \mu)^2 f(x)dx = \int_0^1 \left(x - \frac{1}{2} \right)^2 dx = \left[\frac{1}{3} \left(x - \frac{1}{2} \right)^3 \right]_0^1 = \frac{1}{12}$$

確率変数の変換公式

期待値

確率変数 X と定数 a, b に対して $E\{aX + b\} =$

$Var\{aX + b\} =$

分散

【例題】 期待値 m 標準偏差 σ の確率変数 X を

変換して得られる確率変数

$$Z = \frac{X - m}{\sigma}$$

の期待値と標準偏差を求めよ。

確率変数の変換公式

確率変数 X と定数 a, b に対して $E\{aX + b\} = aE\{X\} + b$

期待値

$$Var\{aX + b\} = a^2 Var\{X\}$$

分散

【例題】 期待値 m 標準偏差 σ の確率変数 X を

変換して得られる確率変数 $Z = \frac{X - m}{\sigma}$ の期待値と標準偏差を求めよ。

$$Z = \frac{X - m}{\sigma}$$

確率変数 X の標準化

$$E\{Z\} = E\left\{\frac{1}{\sigma}X - \frac{m}{\sigma}\right\} = \frac{1}{\sigma}E\{X\} - \frac{m}{\sigma} = \frac{m}{\sigma} - \frac{m}{\sigma} = 0 \quad \text{期待値 } 0$$

$$Var\{Z\} = \frac{1}{\sigma^2}Var\{X\} = \frac{1}{\sigma^2}\sigma^2 = 1 \quad \text{標準偏差 } \sqrt{Var\{Z\}} = 1$$

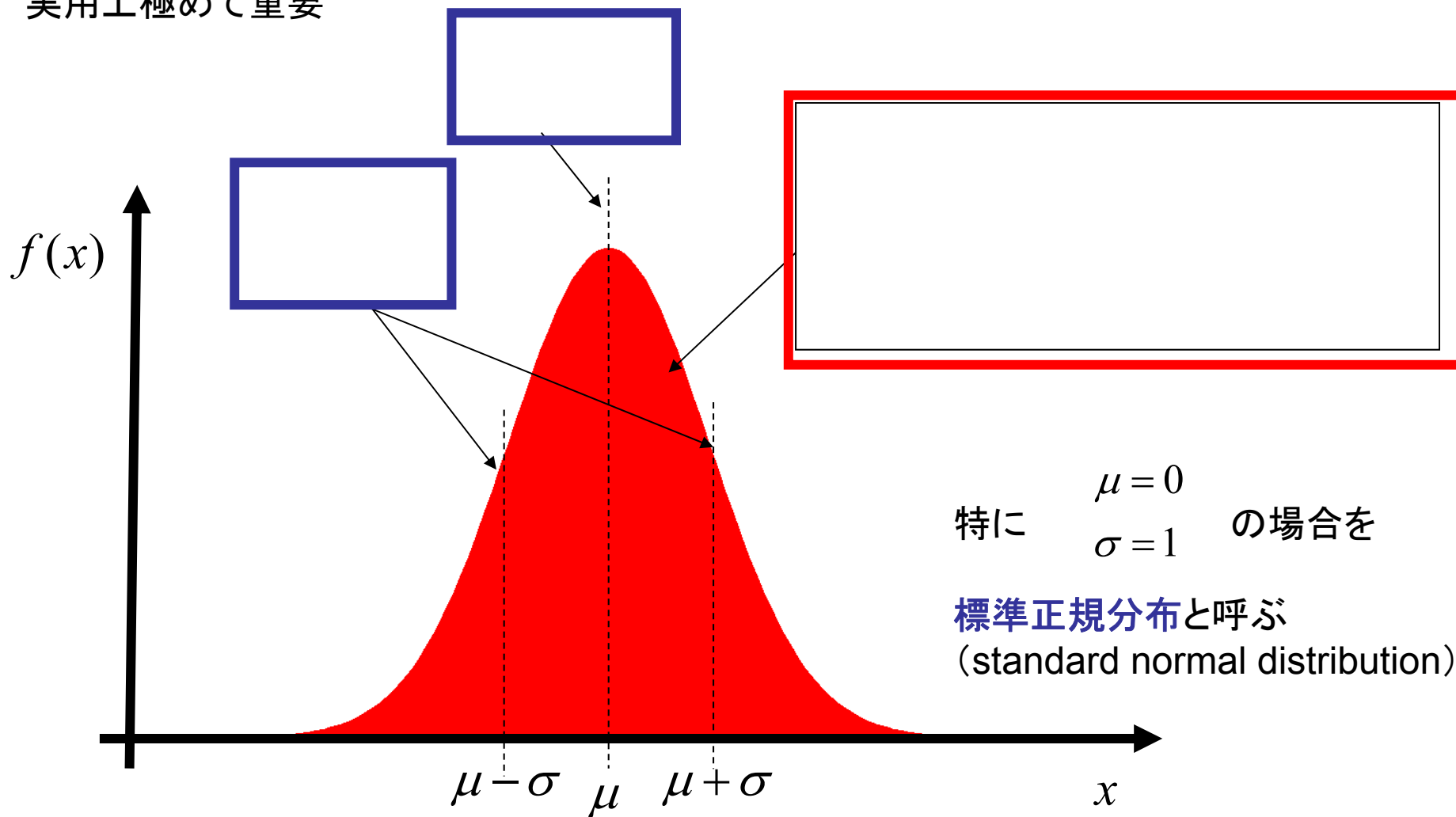
代表的な連続形分布関数：正規分布

Gaussian distribution, normal distribution

$$N(\mu, \sigma^2)$$

期待値 μ
分散 σ^2

多数の分布の和として出てくるような現象の分布：
実用上極めて重要



特に $\mu = 0$ の場合を
 $\sigma = 1$

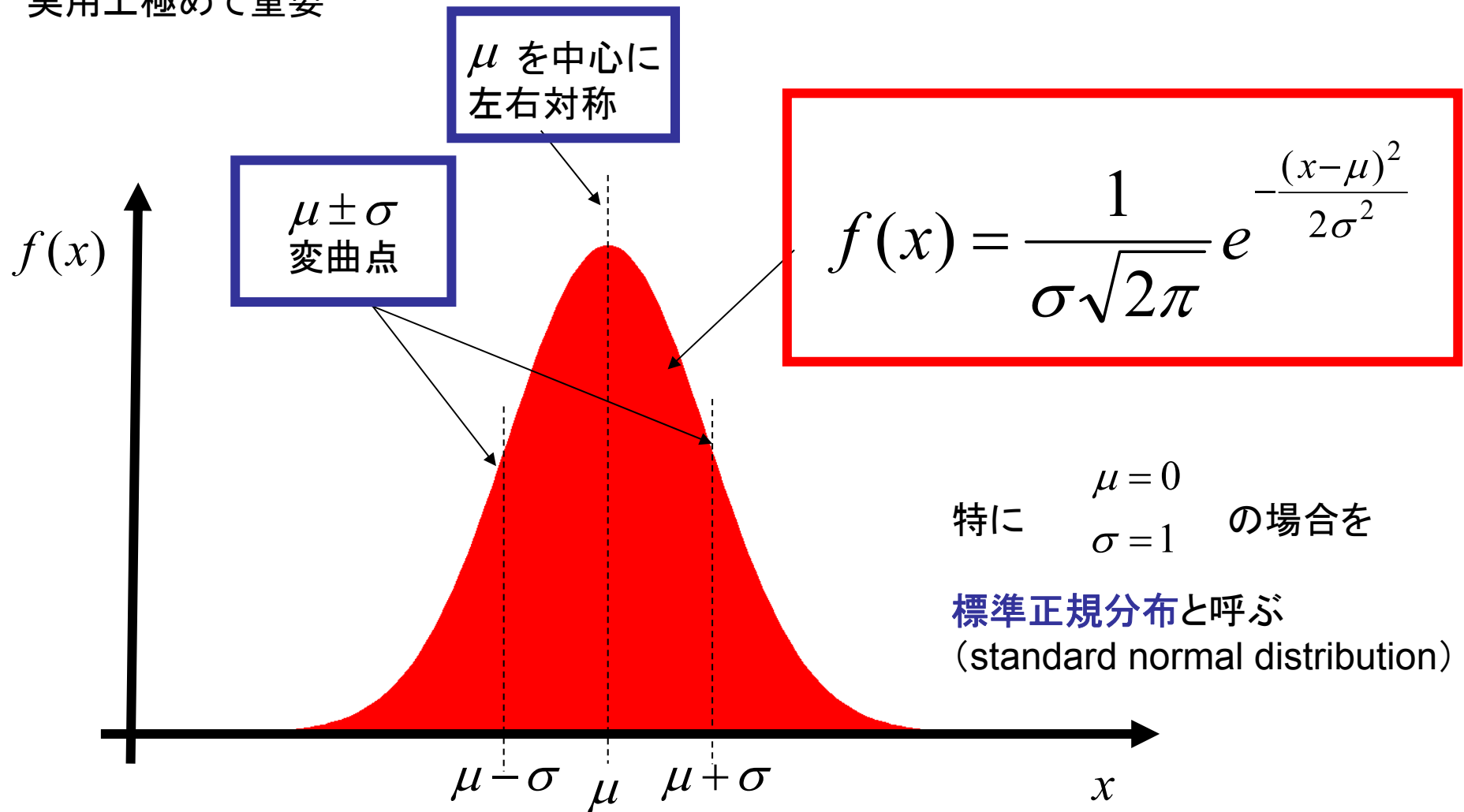
標準正規分布と呼ぶ
(standard normal distribution)

代表的な連続形分布関数：正規分布

Gaussian distribution, normal distribution

$$N(\mu, \sigma^2) \quad \begin{array}{l} \text{期待値 } \mu \\ \text{分散 } \sigma^2 \end{array}$$

多数の分布の和として出てくるような現象の分布：
実用上極めて重要



Johann Carl Friedrich Gauss (1777–1855)

ドイツの数学者・天文学者・物理学者

素数定理の予想

定規とコンパスによる正17角形の作図

代数学(ガウス・ザイデル法)

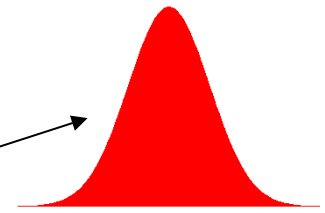
複素関数論(ガウス平面)

最小2乗法、正規分布(ガウス分布)

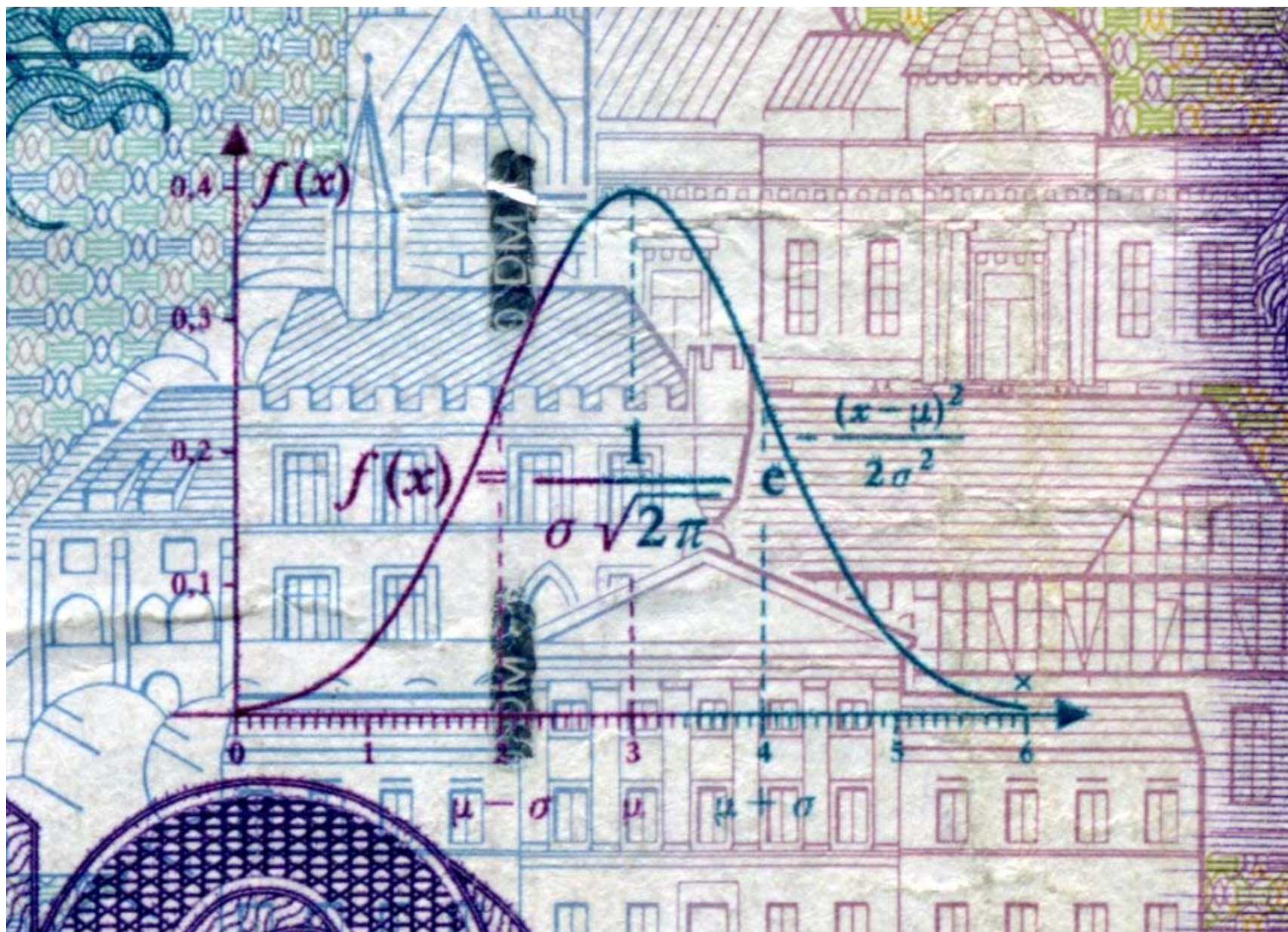
天体望遠鏡(ガウス式レンズ)

微分幾何学・曲面論(ガウス曲率)

電磁気学(ガウスの法則・ガウス(磁束密度の単位))



ドイツの10マルク札 (1999年)
(現在はユーロに統合されたため
流通していない)

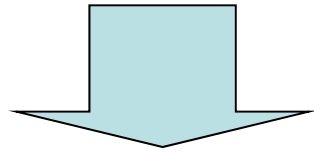


お札にはガウス分布の数式とグラフが！

正規分布における計算

確率密度関数 $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

の定積分 $\int_a^b f(x)dx$ は初等関数で表せない

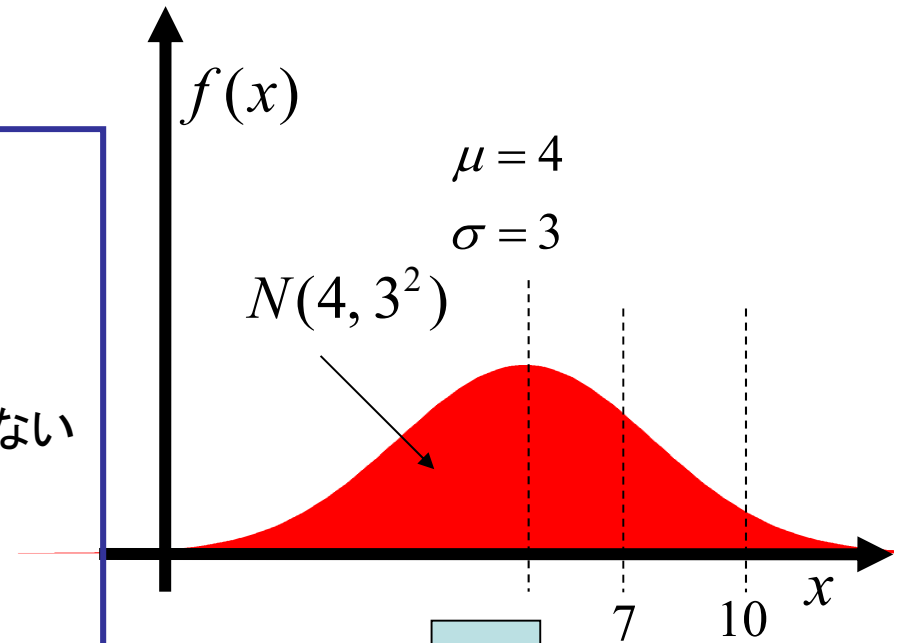


確率変数 x を標準化した上で、
標準正規分布表を用いて計算する

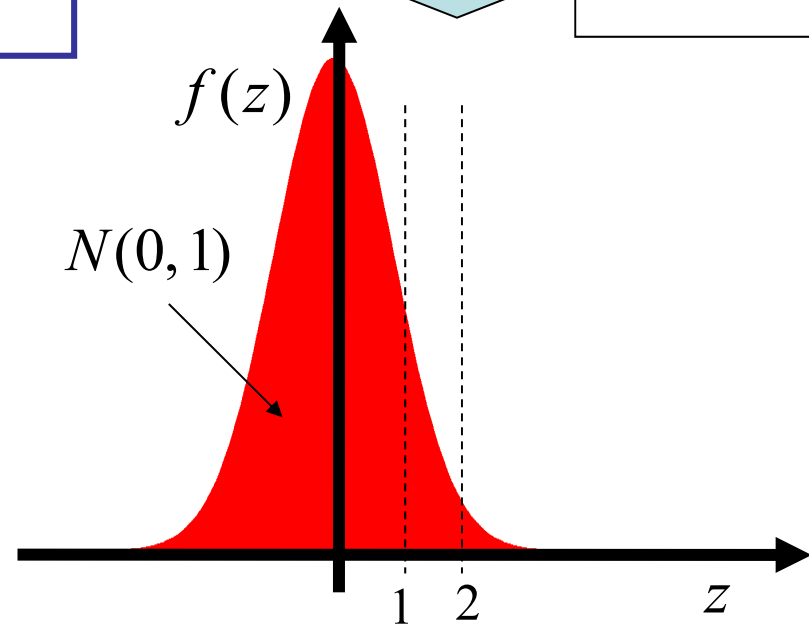
例) $f(x) = N(4, 3^2)$ のとき、

$$P(7 \leq x \leq 10) = \int_7^{10} f(x)dx \text{ を求める}$$

確率変数 $Z = \square$ は $N(0, 1)$ に従う



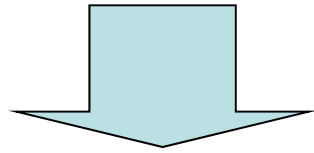
標準化



正規分布における計算

確率密度関数 $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

の定積分 $\int_a^b f(x)dx$ は初等関数で表せない



確率変数 x を標準化した上で、
標準正規分布表を用いて計算する

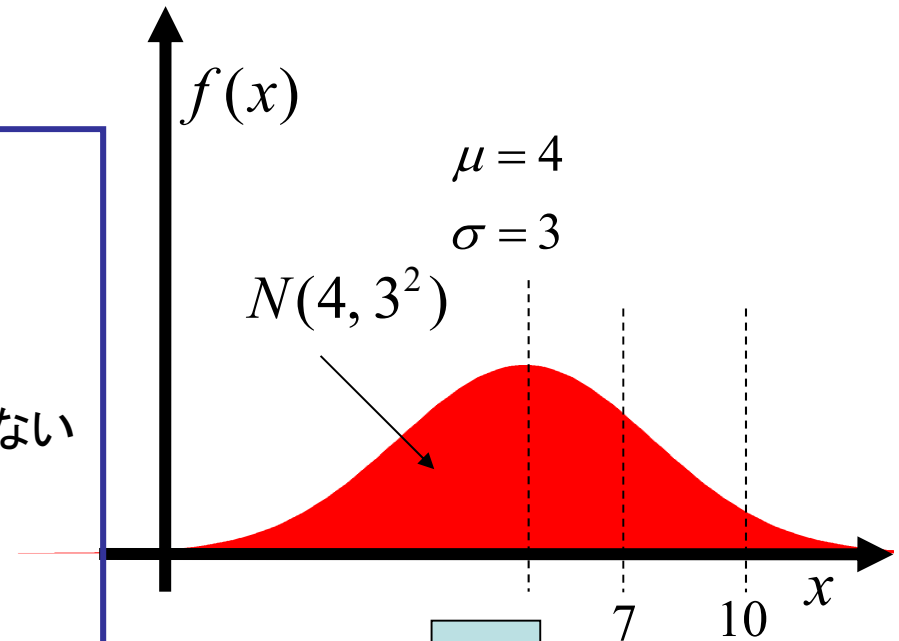
例) $f(x) = N(4, 3^2)$ のとき、

$$P(7 \leq x \leq 10) = \int_7^{10} f(x)dx \text{ を求める}$$

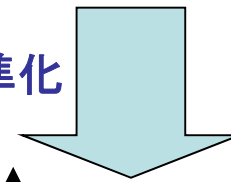
確率変数 $Z = \frac{x-4}{3}$ は $N(0, 1)$ に従う

$$P(7 \leq x \leq 10) = P\left(\frac{7-4}{3} \leq z \leq \frac{10-4}{3}\right)$$

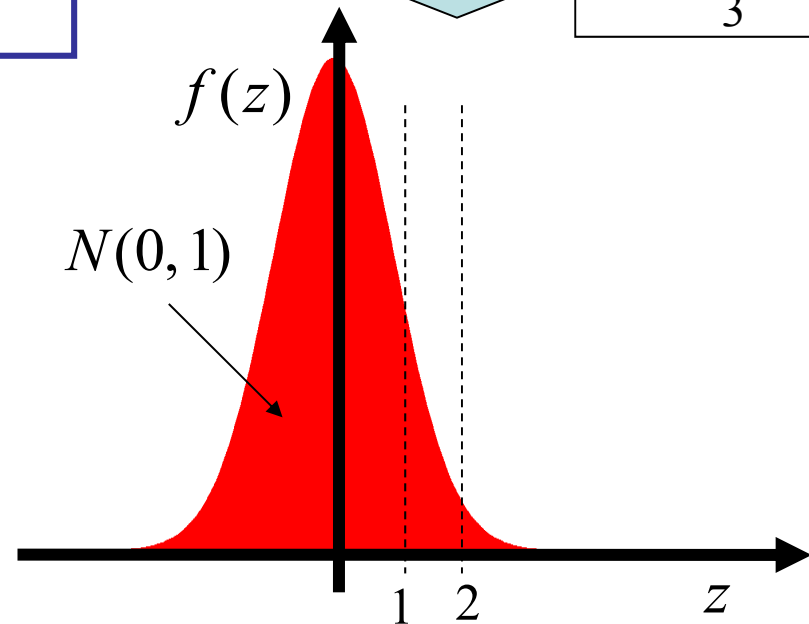
$$= P(1 \leq z \leq 2) \quad \leftarrow \text{あとは表を使って計算}$$



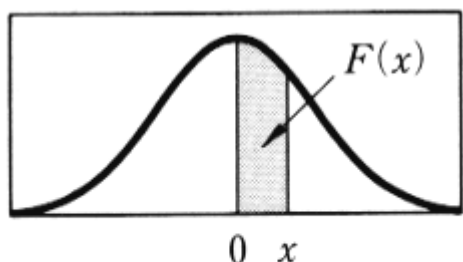
標準化



$$Z = \frac{x-4}{3}$$



標準正規分布表



カゲの部分の確率 $F(x)$ を示す。

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998

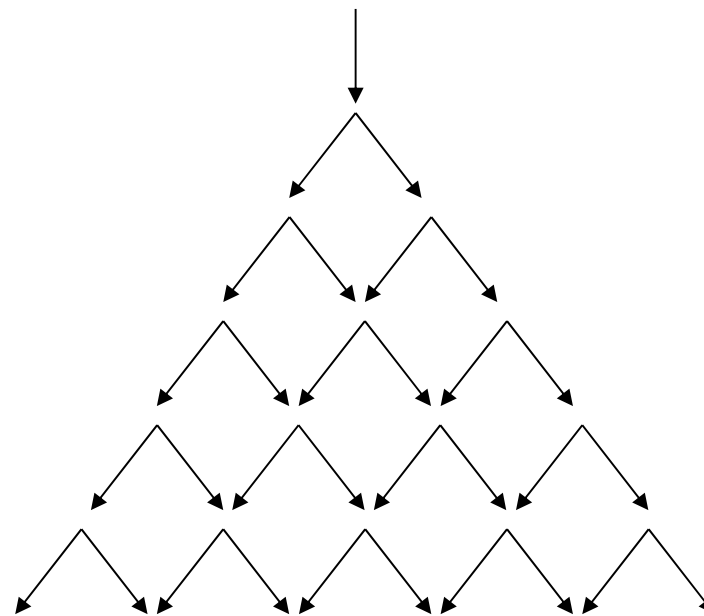
$$\begin{aligned}
 P(7 \leq x \leq 10) &= P\left(\frac{7-4}{3} \leq z \leq \frac{10-4}{3}\right) \\
 &= P(1 \leq z \leq 2) \\
 &= 0.4772 - 0.3413 \\
 &= 0.1359
 \end{aligned}$$

二項分布による正規分布の近似

二項分布 $B(n, p)$ は n が大きくなるとき、
正規分布 に近づく



各段において、右または左へ
 $1/2$ の確率で落ちていく



どこに落ちるかは二項分布に従う

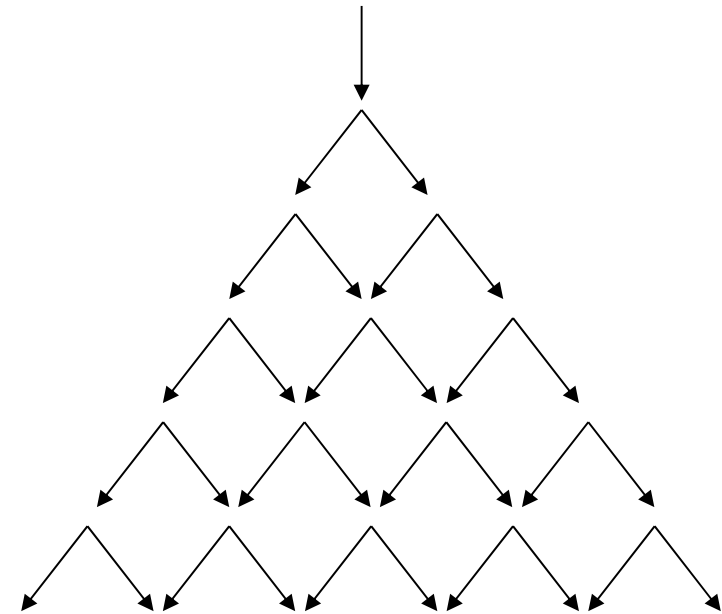
独立なベルヌーイ分布の確率変数の足し合わせ = 二項分布  正規分布へ

二項分布による正規分布の近似

二項分布 $B(n, p)$ は n が大きくなるとき、
正規分布 $N(np, np(1-p))$ に近づく



各段において、右または左へ
 $1/2$ の確率で落ちていく



どこに落ちるかは二項分布に従う

独立なベルヌーイ分布の確率変数の足し合わせ = 二項分布



正規分布へ

【復習】 二項分布

1回の試行で、ある事象の起こる確率が p
この試行を n 回行う

ベルヌイ試行
ベルヌイ分布

確率変数 x : 事象の起こった回数
確率分布 $P(x)$ → 二項分布 $B(n, p)$

(例) 3回サイコロを投げたときの1の目が出る回数

x	0	1	2	3
$P(x)$	$\left(\frac{5}{6}\right)^3 = \frac{125}{216}$	$3\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^2 = \frac{75}{216}$	$3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right) = \frac{15}{216}$	$\left(\frac{1}{6}\right)^3 = \frac{1}{216}$

一般に、確率 p を持つ事象が n 回の試行中 x 回起こることを考えると、

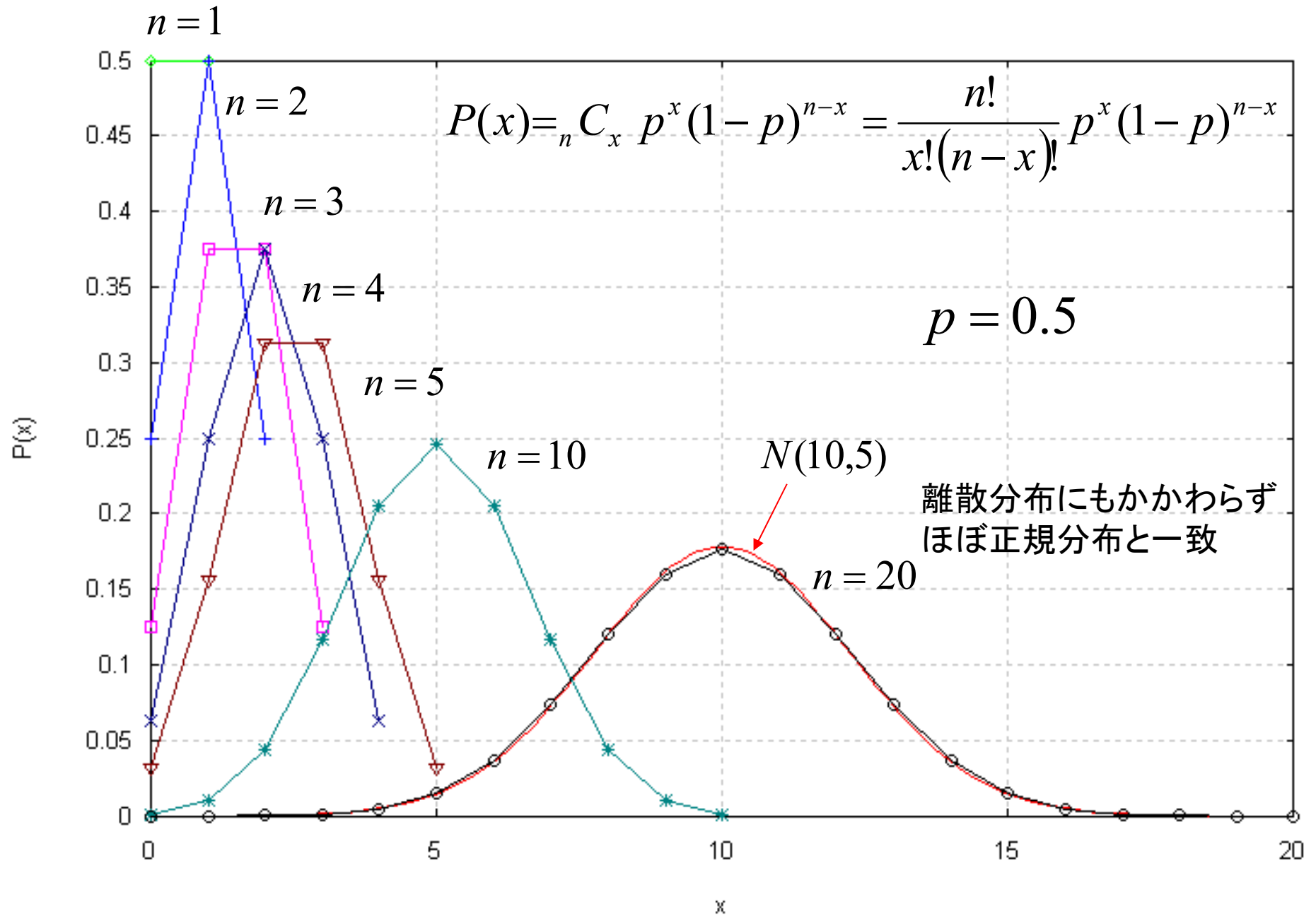
このときの確率は $p^x (1-p)^{n-x}$

$\underbrace{\text{○ ○ … ○}}_{x \text{ 回}} \quad \underbrace{\text{× × … ×}}_{n-x \text{ 回}}$

n 回の試行中に x 回起こるパターンの組合せ (combination) は、

$${}_n C_x = \frac{n!}{x!(n-x)!} \quad \text{よって} \quad P(x) = {}_n C_x p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

二項分布のグラフ



【復習】 二項分布の極限

ポアソン分布(Poisson distribution) 離散分布

単位時間あたり平均で λ 回発生する事象が、
単位時間に x 回(ゼロを含む)発生する確率

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

二項分布において、

確率 p を持つ事象が n 回の試行中 x 回起こることを考えた

→ $\lambda = n p$ とおく

$$P(x) = {}_n C_x p^x (1-p)^{n-x} = \frac{n(n-1)\cdots(n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{-x}$$

ここで

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}, \quad \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) = 1, \quad \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-x} = 1,$$

よって

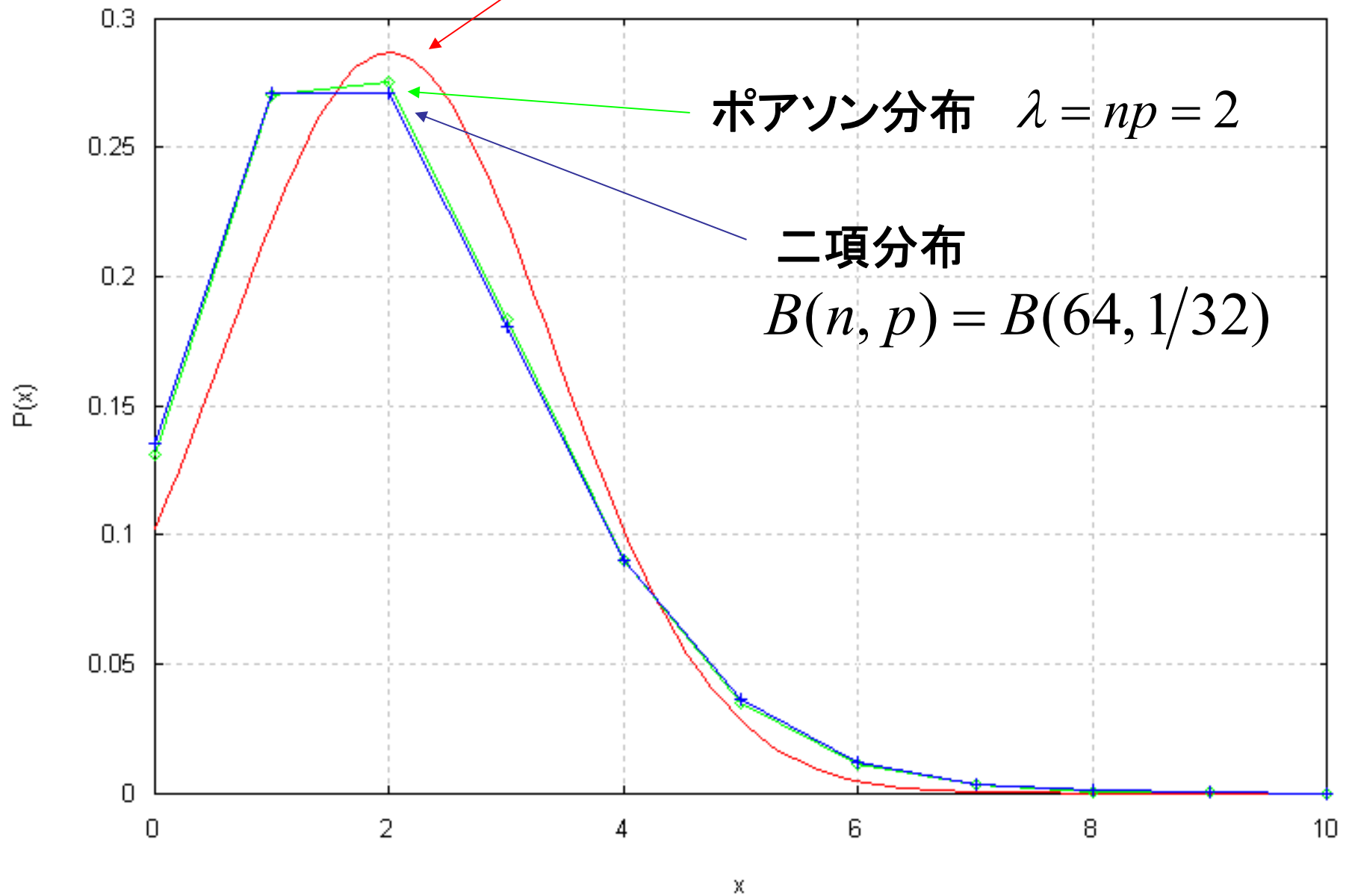
$$P(x) = \lim_{n \rightarrow \infty} {}_n C_x p^x (1-p)^{n-x} = \frac{e^{-\lambda} \lambda^x}{x!}$$

ポアソン分布は
二項分布の極限

正規分布 $N(np, np(1-p)) = N(2, \frac{31}{16})$

ポアソン分布 $\lambda = np = 2$

二項分布
 $B(n, p) = B(64, 1/32)$



まとめ

・確率変数 X が連続値をとる場合: 確率密度関数 $f(x)$

・確率変数 X が $a \sim b$ までの値をとる確率: $P(a \leq x \leq b) = \int_a^b f(x) dx$

・確率変数 X の**期待値** $E\{x\} = \int_{-\infty}^{\infty} x f(x) dx$

・確率変数 X の**分散** $Var\{x\} = E\{(x - E\{x\})^2\} = \int_{-\infty}^{\infty} (x - E\{x\})^2 f(x) dx$

・期待値 m 標準偏差 σ の確率変数 X の標準化 $Z = \frac{X - m}{\sigma}$

期待値 0
分散 1

・代表的な確率密度関数の例

1) **一様分布** (コンピュータで簡単に生成可能)

2) **指数分布** (機械が偶発的に故障するまでの時間 / 客が訪れる時間間隔等)

ポアソン分布でも表される同一現象の異なる表現

3) **正規分布** (さまざまな乱数の合計の極限)

【演習問題】 2018.04.24

学籍番号

氏名

指数分布 $f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$

【ヒント】 部分積分の公式

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$$

の期待値および分散を求めよ

【期待値】

【分散】

【演習問題】

学籍番号

氏名

指数分布 $f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$ **【ヒント】** 部分積分の公式

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$$

の期待値および分散を求めよ

$$\begin{aligned} E\{x\} &= \int_0^{\infty} x \lambda e^{-\lambda x} dx = \left[-x\lambda e^{-\lambda x} \right]_0^{\infty} - \int_0^{\infty} -e^{-\lambda x} dx \\ &= \frac{1}{\lambda} \end{aligned}$$

$$\begin{aligned} V\{x\} &= \int_0^{\infty} \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x} dx = \left[-\left(x - \frac{1}{\lambda}\right)^2 e^{-\lambda x} \right]_0^{\infty} - \int_0^{\infty} -2\left(x - \frac{1}{\lambda}\right) e^{-\lambda x} dx \\ &= \frac{1}{\lambda^2} \end{aligned}$$