

## 確率論 (Probability Theory) 第6週

tv.yamamoto

### 1 累積分布関数

標本空間を  $X$  と書く。 $X$  が離散集合であれば、その上の確率変数は離散確率変数、確率分布は離散分布と呼ばれる。 $X$  が連続集合であれば、その上の確率変数は連続確率変数、確率分布は連続確率分布と呼ばれる。

まず離散確率分布を考える。離散確率事象の標本空間  $X$  はしばしば整数の集合で表現される。例えば、コイン投げでは表が 1、裏が 0 であり  $X = \{0, 1\}$ 、賽子投げでは出る目の数字そのものであり  $X = \{1, 2, 3, 4, 5, 6\}$  となる。その要素を  $k$  と記す。そのような集合が標本空間であれば、 $\{k\}, k \in X$  は標本空間のひとつの分割を与える。

$$X = \{0, 1, 2, 3, 4, \dots\} = \{0\} \cup \{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \dots \quad (1)$$

よって  $\forall k \in X$  について、確率の公理や、加法法則、乗加法則などを満たすように、確率

$$0 \leq Pr[X = k] \leq 1 \quad (2)$$

を定義できる。これを確率関数と呼び、

$$p(k) = Pr[X = k] \quad (3)$$

と書く。

次に連続確率分布を考える。連続確率事象の標本空間  $X$  はしばしば実数の集合で表現される。例えば、人間の身長や体重は正の実数の一部、人間や製品の寿命は非負の実数の一部、実験で体験する測定誤差や実験誤差は実数全体、である。その要素を  $x$  と記すことにする。それら実数空間は連続集合であり、標本空間の分割は点集合に基づくのではなく、次のような半開区間に基づいて構成する。任意の単調実数列  $\inf X = a_0 < a_1 < a_2 < \dots < a_n < \dots$  は標本空間のひとつの分割

$$(a_0, a_1], (a_1, a_2], (a_2, a_3], \dots, (a_{n-1}, a_n], (a_n, a_{n+1}], \dots \quad (4)$$

を与える。これらが互いに疎であり、すべての和集合が標本空間  $X$  を覆うことは、定義より明らか。連続確率分布は、それぞれの半開区間について確率

$$0 \leq Pr[X \in (a_k, a_{k+1}]] \leq 1 \quad (5)$$

を与える。この与え方は、分割を変えても良いので、確率関数として、

$$F(a, b) = Pr[X \in (a, b]] \quad (6)$$

を定めることが考えられる。ただし、引数が二つあり、数学的な取り扱いが不便となるため、連続確率分布の確率関数を次のように定義する。

先の分割に基づいて、上の部分集合を左から順に和集合を作ると

$$(a_0, a_1], (a_0, a_2], (a_0, a_3], \dots, (a_0, a_n], (a_0, a_{n+1}], \dots \quad (7)$$

という部分集合の列を得る。これらは互いに疎ではないが、

$$(a_0, a_1] \subset (a_0, a_2] \subset (a_0, a_3] \subset \dots \subset (a_0, a_n] \subset (a_0, a_{n+1}] \subset \dots \quad (8)$$

を満たす、単調増加な集合の列という性質を持つ。このような集合列の全体に対して確率を

$$F(x) = Pr[X \leq x] = Pr[X \in (-\infty, x]] \quad (9)$$

と定める。これを 累積分布関数 と言う。上の集合列は、互いに疎な区間の和集合として表されていることに注意する。

離散確率分布についても同様に

$$F(x) = Pr[X \leq x] = \sum_{k \leq x, k \in \mathcal{X}} p(k) \quad (10)$$

と累積分布関数を定めることができる。

確率分布の累積分布関数が  $F(x)$  あるいは  $F(k)$  と与えられれば、確率変数  $X$  が区間  $(a, b]$  に含まれる確率は

$$Pr[X \in (a, b]] = Pr[X \in (-\infty, b]] - Pr[X \in (-\infty, a]] = F(b) - F(a) \quad (11)$$

と求めることができる。

話を連続確率分布に戻す。連続確率分布の累積分布関数が、 $x$  に関して微分可能なとき、その微分

$$f(x) = \frac{d}{dx} F(x) \quad (12)$$

を、確率分布  $F$  の確率密度関数 (あるいは単に密度関数) と呼ぶ。確率密度関数は、確率変数  $X$  が微小区間  $(x - \Delta x/2, x + \Delta x/2]$  に含まれる確率が

$$Pr[X \in (x - \Delta x/2, x + \Delta x/2]] \simeq f(x) \Delta x \quad (13)$$

と近似できることから、 $f(x)$  は点  $x$  の近傍での確率の大小を意味することが分かる。

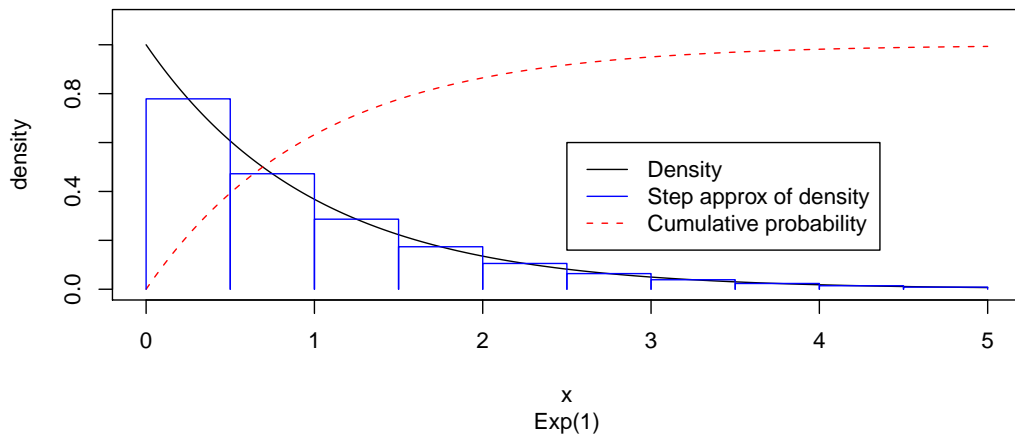


図 1  $\lambda = 1$  の指数分布の確率密度関数とその階段近似

この話をもう少し進めると、標本空間のひとつの分割

$$(a_0, a_1], (a_1, a_2], (a_2, a_3], \dots, (a_{n-1}, a_n], (a_n, a_{n+1}], \dots \quad (14)$$

のそれぞれの区間に対して、区間幅を  $\Delta_k$ 、幅確率を

$$Pr[X \in (a_k, a_{k+1}]] = p(a_k) \quad (15)$$

と置く。すると

$$(a_k, a_{k+1}] \cap (a, a_2] = \{\phi\} \quad (16)$$

であり、

$$Pr[X \in \mathcal{X}] = \sum_{k=0}^{\infty} p(a_k) \quad (17)$$

今、区間幅を  $\Delta_k \equiv \Delta$  に統一すると、

$$Pr[X \in (a_k, a_{k+1}]] \simeq f\left(\frac{a_k + a_{k+1}}{2}\right)\Delta \quad (18)$$

となる。このような階段近似の一例が、図 1 の階段関数である。

もし標本空間が 2 次元以上のユークリッド空間であれば、その次元を  $p$  と置くと、累積分布関数は

$$F(x_1, x_2, \dots, x_p) = Pr[X_1 \leq x_1 \text{ and } X_2 \leq x_2 \text{ and } \dots \text{ and } X_p \leq x_p] \quad (19)$$

またその確率密度関数は

$$f(x_1, x_2, \dots, x_p) = \frac{d^p}{dx_1 dx_2 \dots dx_p} F(x_1, x_2, \dots, x_p) \quad (20)$$

と定める。

## 2 期待値

確率分布  $F$  に関する期待値演算は  $E_F[\cdot]$  と書く。括弧の中身の  $F$  に関する期待値を計算する。例えば  $F$  に従う確率変数  $X$  の期待値は

$$E[X] = \begin{cases} \sum_{k \in \mathcal{X}} k p(k) & F \text{ が離散分布の場合} \\ \int_{x \in \mathcal{X}} x f(x) dx & F \text{ が連続分布の場合} \end{cases} \quad (21)$$

となる。期待値は確率変数  $X$  のそれとは限らず、 $X^2$  や  $(X - E[X])^2$  など、確率変数  $X$  の関数の期待値を計算することもあり、 $X$  の関数  $g(X)$  の期待値を、離散分布と連続分布それぞれについて

$$E[g(X)] = \begin{cases} \sum_{k \in \mathcal{X}} g(k) p(k) & F \text{ が離散分布の場合} \\ \int_{x \in \mathcal{X}} g(x) f(x) dx & F \text{ が連続分布の場合} \end{cases} \quad (22)$$

と定める。例えば離散分布の場合に、標本空間が非負の整数全体の集合であれば、

$$E[g(X)] = \sum_{k \in \mathcal{X}} g(k) p(k) = \sum_{k=0}^{\infty} g(k) p(k) \quad (23)$$

また連続分布の場合に、標本空間が実数全体の集合であれば、

$$E[g(X)] = \int_{x \in \mathcal{X}} g(x) f(x) dx = \int_{-\infty}^{\infty} g(x) f(x) dx \quad (24)$$

あるいは非負の実数全体の集合であれば、

$$E[g(X)] = \int_{x \in \mathcal{X}} g(x) f(x) dx = \int_0^{\infty} g(x) f(x) dx \quad (25)$$

などと等しくなる。

$$E[g(X)] = \sum_{k \in \mathcal{X}} g(k) p(k) = \sum_{k=0}^{\infty} g(k) p(k) \quad (26)$$

前節で示し、図 1 に例示した階段関数近似を用いて、連続確率分布に関する  $g(X)$  の期待値を説明することもできる。連続確率変数  $X$  の代わりに、前節で導入した標本空間  $X$  の分割のそれぞれの区間の番号を選ぶ離散確率変数  $Y$  を定める。

$$Y(X) = k \text{ if } X \in (a_k, a_{k+1}] \quad (27)$$

このような  $Y$  は  $X$  の関数であり、非負の整数のみをとる。 $Y$  の確率関数は

$$p(k) = Pr[Y = k] = Pr[X \in (a_k, a_{k+1}]] \quad (28)$$

である。この  $Y$  に関して、更に  $h(Y) = (a_Y + a_{Y+1})/2$  という関数を考える。これは

$$E[h(Y)] = \sum_{k=0}^{\infty} h(k) p(k) = \sum_{k=0}^{\infty} \frac{a_k + a_{k+1}}{2} f\left(\frac{a_k + a_{k+1}}{2}\right) (a_{k+1} - a_k) \quad (29)$$

である。右辺の分割の数を無限大にしたときの極限が存在すれば、それが積分

$$\int_X x f(x) dx = E[X] \quad (30)$$

であったことを思い起こすと、離散確率分布と連続確率分布の関係が少しは繋がって見えてくるかもしれない。よく用いられる期待値には、次のようなものがある。

- $E[X]$ : (無限回の試行を繰り返した時に) 観測される確率変数が平均して幾つぐらいの値をとるか。 $X$  の期待値、 $F$  の期待値、あるいは単に 期待値 と呼ばれ、しばしば  $\mu$ 、 $\mu_1$ 、あるいは  $m_1$  と記される。
- $E[X^2]$ : (無限回の試行を繰り返した時に) 観測される確率変数の二乗値が、平均して幾つぐらいの値をとるか。 $X^2$  の二乗の期待値 と呼ばれ、たまに  $m_2$  と記される。
- $V[X] = E[(X - E[X])^2]$ : (無限回の試行を繰り返した時に) 観測される確率変数  $X$  は、その期待値  $E[X]$  から、平均して、二乗距離<sup>1</sup>でどのぐらい離れているか。これは 分散 と呼ばれ、しばしば  $\sigma^2$  あるいは  $\mu_2$  と表される。また、分散の平方根  $\sigma$  あるいは  $\sqrt{\mu_2}$  を 標準偏差 と呼ぶ。分散の単位は  $X$  もしくは  $\mu$  の単位の二乗であり、標準偏差の単位は  $X$  もしくは  $\mu$  と等しくなることに注意する。

これらの期待値が、確率分布の特徴量となる。試みに上の図 1 に、平均値の線と確率分布の平均と分散が分かれば、確率関数の大凡の中心とその回りでの広がり具合が想像できることから分かる。

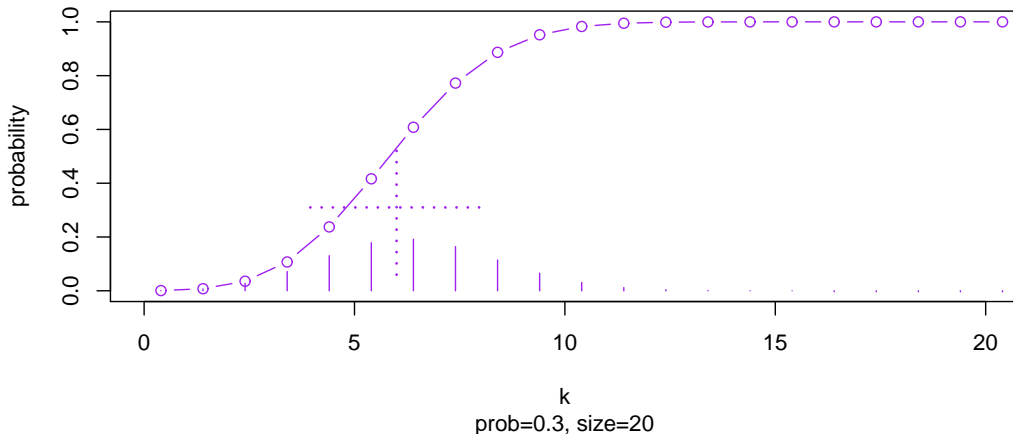


図 2  $n = 20, p = 0.3$  の二項分布の確率関数と期待値と標準偏差

<sup>1</sup> $a$  と  $b$  の二乗距離は  $d_2(a, b) = (a - b)^2$ 。二乗は非負なので、 $d_2(a, b) = d_2(b, a)$

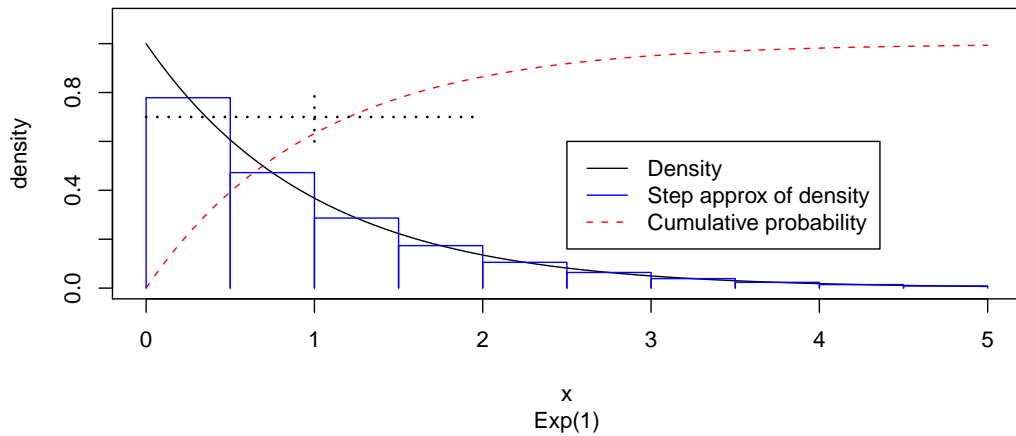


図 3  $\lambda = 1$  の指数分布の確率密度関数と期待値と標準偏差

$X^k$  の期待値  $E[X^k]$  には、 $X$  の原点のまわりの  $k$  次のモーメント、という名前がついている。これは、しばしば次数  $k$  を添え字に取り、

$$m_k = E[X^k] \quad (31)$$

と記される。一方、 $(X - \mu)^k$  の期待値  $E[(X - \mu)^k]$  には、 $X$  の平均のまわりの  $k$  次のモーメント、という名前がついている。これは、しばしば次数  $k$  を添え字に取り、

$$\mu_k = E[(X - \mu)^k] \quad (32)$$

と記される。

期待値に関して、幾つかの関係式がよく用いられる。

$$\begin{aligned}
 E[aX + b] &= \int_{x \in X} (ax + b) f(x) dx \\
 &= \int_{x \in X} axf(x) dx + \int_{x \in X} bf(x) dx \\
 &= a \int_{x \in X} xf(x) dx + b \int_{x \in X} f(x) dx \\
 &= aE[X] + b \\
 E[aX + b] &= \sum_{k \in X} (ak + b) p(k) \\
 &= \sum_{k \in X} akp(k) + \sum_{k \in X} bp(k) \\
 &= a \sum_{k \in X} kp(k) + b \sum_{k \in X} p(k) \\
 &= aE[X] + b
 \end{aligned} \quad (33)$$

この関係は例えば、

$$E[aX^2 + b] = aE[X^2] + b \quad (34)$$

など、 $X$  の関数についても

$$E[ag(X) + b] = aE[g(X)] + b \quad (35)$$

のように成り立つ。

分散  $(X - \mu)$  の二乗の期待値) についても、

$$\begin{aligned}
 V[X] &= \sigma^2 = \mu_2 \\
 &= E[(X - \mu)^2] = \int_{x \in \mathcal{X}} (x - \mu)^2 f(x) dx \\
 &= \int_{x \in \mathcal{X}} x^2 f(x) dx - 2 \int_{x \in \mathcal{X}} x\mu f(x) dx + \int_{x \in \mathcal{X}} \mu^2 f(x) dx \\
 &= E[X^2] - 2\mu \int_{x \in \mathcal{X}} x f(x) dx + \mu^2 \int_{x \in \mathcal{X}} f(x) dx \\
 &= E[X^2] - 2\mu E[X] + \mu^2 \\
 &= E[X^2] - (E[X])^2 = m_2 - m_1^2
 \end{aligned} \tag{36}$$

$$\begin{aligned}
 V[X] &= \sigma^2 = \mu_2 \\
 &= E[(X - \mu)^2] = \sum_{x \in \mathcal{X}} (x - \mu)^2 p(x) \\
 &= \sum_{x \in \mathcal{X}} x^2 p(x) - 2 \sum_{x \in \mathcal{X}} x\mu p(x) + \sum_{x \in \mathcal{X}} \mu^2 p(x) \\
 &= E[X^2] - 2\mu \sum_{x \in \mathcal{X}} xp(x) + \mu^2 \sum_{x \in \mathcal{X}} p(x) \\
 &= E[X^2] - 2\mu E[X] + \mu^2 \\
 &= E[X^2] - (E[X])^2 = m_2 - m_1^2
 \end{aligned} \tag{37}$$

および、任意の定数  $a$  について

$$\begin{aligned}
 V[aX] &= E[(aX - a\mu)^2] = \int_{x \in \mathcal{X}} (ax - a\mu)^2 f(x) dx \\
 &= \int_{x \in \mathcal{X}} a^2 (x - \mu)^2 f(x) dx \\
 &= a^2 \int_{x \in \mathcal{X}} (x - \mu)^2 f(x) dx \\
 &= a^2 V[X] = a^2 \sigma^2
 \end{aligned} \tag{38}$$

$$\begin{aligned}
 V[aX] &= E[(aX - a\mu)^2] = \sum_{x \in \mathcal{X}} (ax - a\mu)^2 f(x) dx \\
 &= \sum_{x \in \mathcal{X}} a^2 (x - \mu)^2 f(x) dx \\
 &= a^2 \sum_{x \in \mathcal{X}} (x - \mu)^2 f(x) dx \\
 &= a^2 V[X] = a^2 \sigma^2
 \end{aligned} \tag{39}$$

が成り立つ。前者の関係は「分散は二乗の平均引く平均の二乗」、後者の関係は「定数倍の分散は分散の定数の二乗倍」と覚える。

### 3 二項分布

1回の試行で、成功する確率を  $p$ 、失敗する確率を  $1 - p$  とする。これをベルヌーイ試行と呼び、その確率関数は

$$p(0) = Pr[X = 0] = 1 - p, \quad p(1) = Pr[X = 1] = p \tag{40}$$

となる。毎回の試行を独立<sup>2</sup>とすると、この試行を 2 回繰り返したとき、成功する回数は {0, 1, 2} の何れかとなる。それぞれの値を取る確率は

$$\begin{aligned} p(0) &= Pr[X_1 = 0 \text{ and } X_2 = 0] \\ &= Pr[X_1 = 0] Pr[X_2 = 0] = (1 - p)^2 \end{aligned} \quad (41)$$

$$\begin{aligned} p(1) &= Pr[\{X_1 = 0 \text{ and } X_2 = 1\} \text{ or } \{X_1 = 1 \text{ and } X_2 = 0\}] \\ &= Pr[X_1 = 1] Pr[X_2 = 0] + Pr[X_1 = 0] Pr[X_2 = 1] \\ &= p(1 - p) + (1 - p)p = 2p(1 - p) \\ &= {}_2C_1 p(1 - p) \end{aligned} \quad (42)$$

$$\begin{aligned} p(2) &= Pr[X_1 = 1 \text{ and } X_2 = 1] \\ &= Pr[X_1 = 1] Pr[X_2 = 1] = p^2 \end{aligned} \quad (43)$$

$n$  回の互いに独立な試行で  $k$  回の成功を観測する確率は、

$$p(k) = Pr[X = k] = {}_n C_k p^k (1 - p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1 - p)^{n-k} \quad (44)$$

となる。この確率関数を持つ確率分布は 二項分布 と呼ばれる。この確率分布の名前の由来は、次のような 二項展開 に依る。

$$\begin{aligned} (p + (1 - p))^n &= 1 \\ &= {}_n C_0 p^0 (1 - p)^{n-0} + {}_n C_1 p^1 (1 - p)^{n-1} + {}_n C_2 p^2 (1 - p)^{n-2} + \dots \\ &\quad + {}_n C_{n-1} p^{n-1} (1 - p)^1 + {}_n C_n p^n (1 - p)^0 \end{aligned} \quad (45)$$

この式は、全確率が 1 であることも意味している。

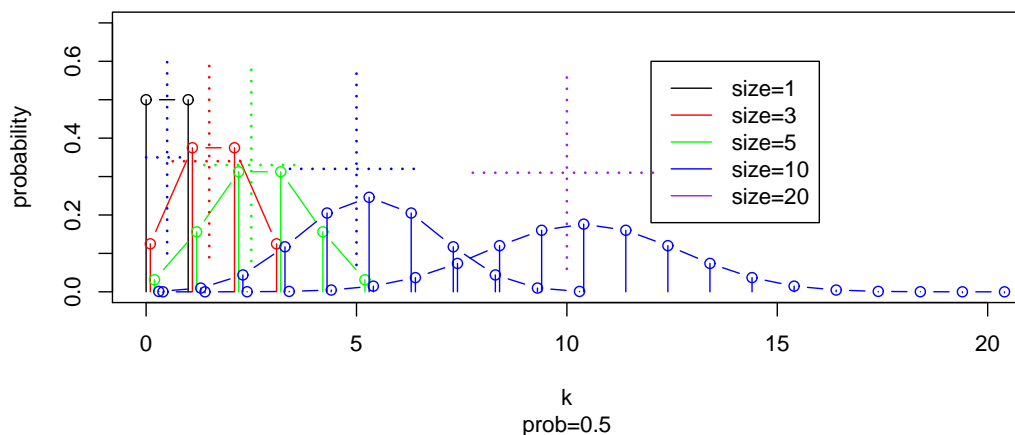
確率関数

$$p(k) = p(k; n, p) = Pr[X = k] = {}_n C_k p^k (1 - p)^{n-k} \quad (46)$$

累積分布関数

$$F(k) = Pr[X \leq k] = \sum_{l=0}^k {}_n C_l p^l (1 - p)^{n-l} \quad (47)$$

ここまでをグラフに表すと、次の通り。まずは確率関数から。



<sup>2</sup>独立ならば同時に起こる確率はそれぞれが単独で起こる確率の積となる。 $Pr[X_1 \leq a \text{ and } X_2 \leq b] = Pr[X_1 \leq a] Pr[X_2 \leq b]$  も、 $Pr[X_1 = a \text{ and } X_2 = b] = Pr[X_1 = a] Pr[X_2 = b]$  も成り立つ。

図4  $p = 0.5$  の二項分布の確率関数 ( $n = 1, 3, 5, 10, 20$ )

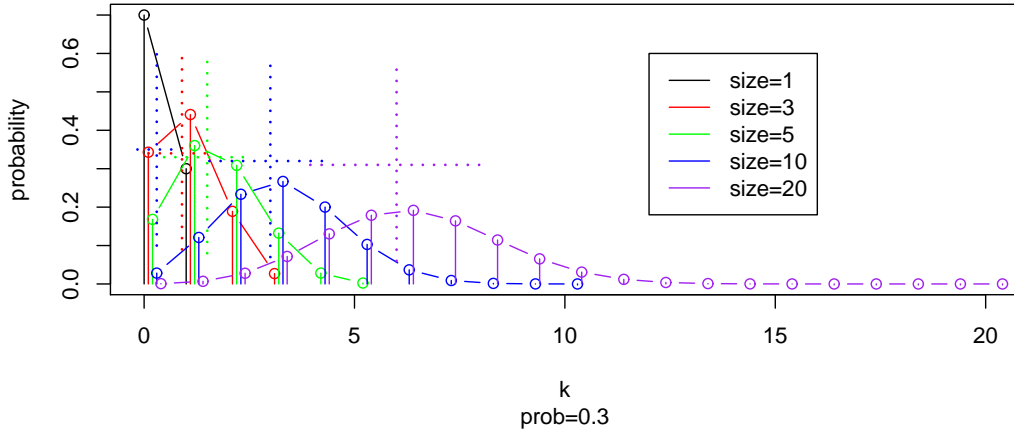


図5  $p = 0.3$  の二項分布の確率関数 ( $n = 1, 3, 5, 10, 20$ )

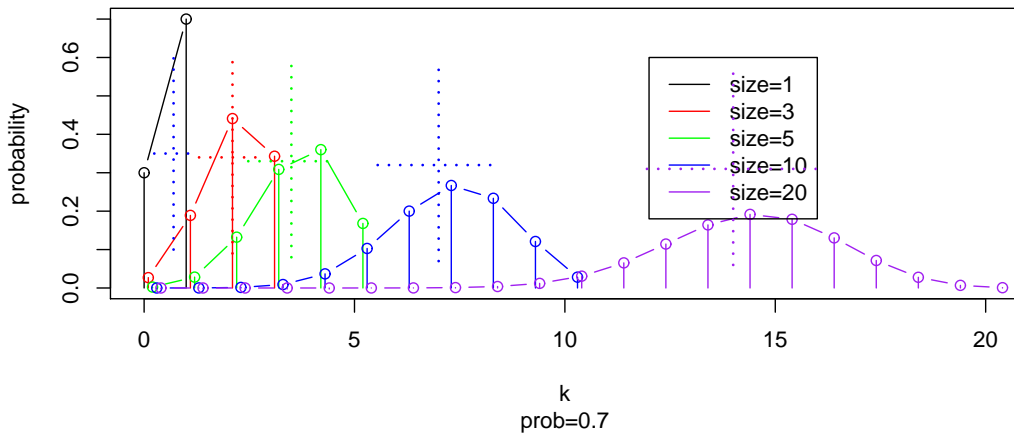


図6  $p = 0.7$  の二項分布の確率関数 ( $n = 1, 3, 5, 10, 20$ )

確率関数は  $p = 0.5$  の時に対象、また  $p < 0.5$  では右に偏り、 $p > 0.5$  では左に偏る。また、 $n$  回の試行を繰り返すと  $np$  に近い整数の方が、 $np$  より遠い整数よりも、観測されやすい。

以下は、確率関数と累積分布関数の関係の例である。総試行回数が 2 の場合は、

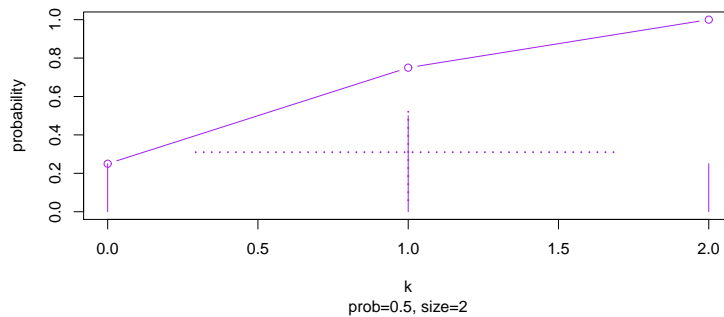




図 7  $n = 2, p = 0.5$  の二項分布の確率関数と累積分布関数

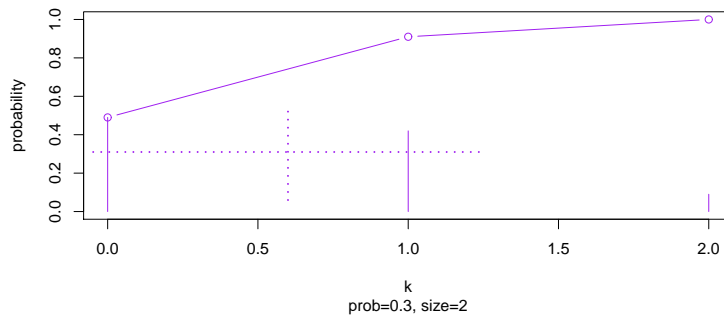


図 8  $n = 2, p = 0.3$  の二項分布の確率関数と累積分布関数

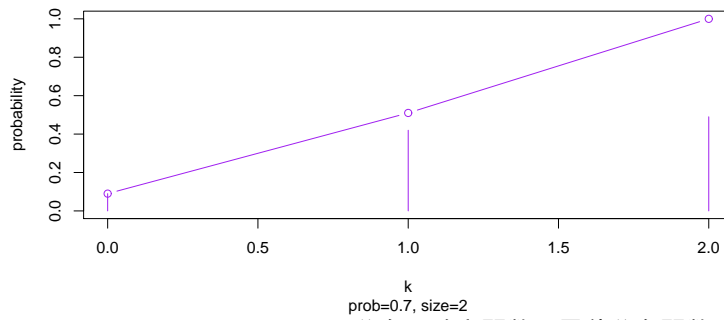


図 9  $n = 2, p = 0.7$  の二項分布の確率関数と累積分布関数

のように、自分で計算して描くことも簡単だろう。総回数が 20 の場合は、

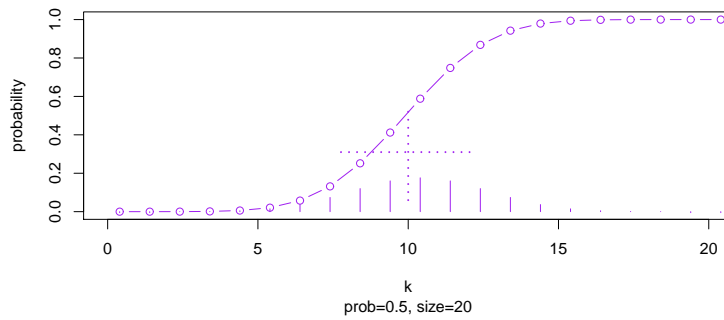


図 10  $n = 20, p = 0.5$  の二項分布の確率関数と累積分布関数

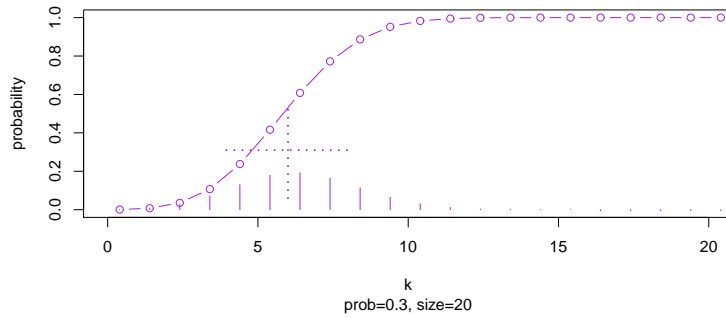


図 11  $n = 20, p = 0.3$  の二項分布の確率関数と累積分布関数

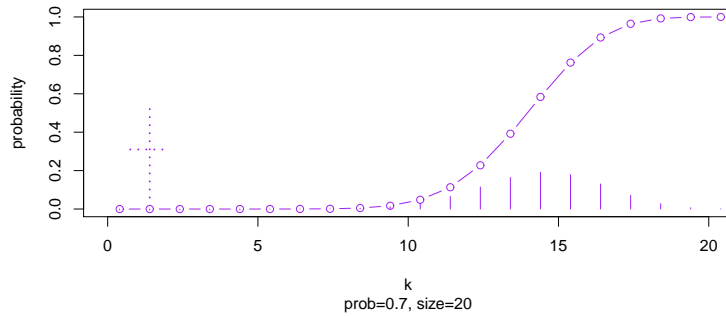


図 12  $n = 20, p = 0.7$  の二項分布の確率関数と累積分布関数

となる。

$X$  の期待値 (=  $X$  の平均)

$$E[X] = \sum_{k=0}^n k {}_n C_k p^k (1-p)^{n-k} = \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} =? \quad (48)$$

$X^2$  の期待値 (=  $X^2$  の平均)

$$E[X^2] = \sum_{k=0}^n k^2 {}_n C_k p^k (1-p)^{n-k} = \sum_{k=0}^n k^2 \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} =? \quad (49)$$

$X$  の分散 (=  $X^2$  の平均 -  $X$  の平均の二乗)

$$V[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 =? \quad (50)$$

## 4 指数分布 #1

指数分布は、ランダムな到着間隔の確率分布、あるいは待ち時間の分布、として用いられることが多い。オペレーションズリサーチで待ち行列というトピックを学ぶ際にも、M/M/1 という基本的なモデルの到着間隔と処理時間の確率分布として紹介される。

確率密度関数

$$f(x) = \frac{d}{dx} (1 - \exp(-\lambda x)) =? \quad (51)$$

累積分布関数

$$F(k) = Pr[X \leq k] = 1 - \exp(-\lambda x) \quad (52)$$

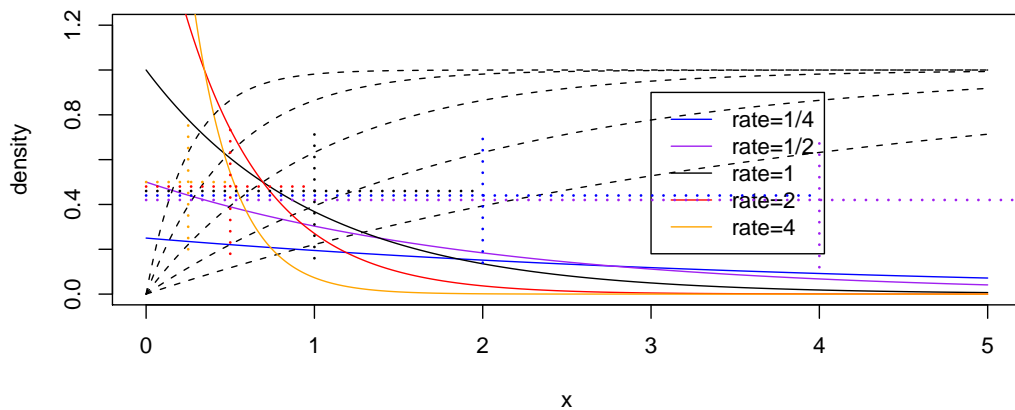


図 13 指数分布の累積分布関数と確率密度関数

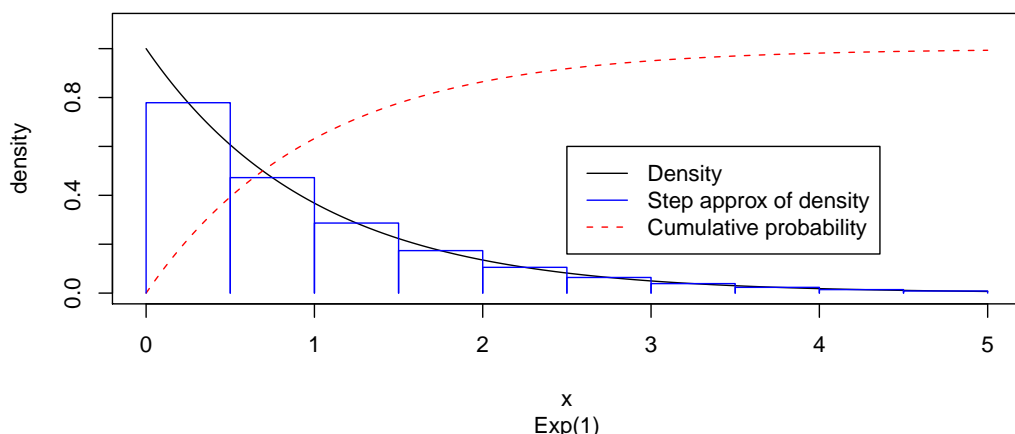


図 14  $\lambda = 1$  の指数分布の累積分布関数と確率密度関数

$X$  の期待値 (=  $X$  の平均)

$$E[X] = \int_0^{\infty} xf(x) = ? \quad (53)$$

$X^2$  の期待値 (=  $X^2$  の平均)

$$E[X^2] = \int_0^{\infty} x^2f(x) = ? \quad (54)$$

$X$  の分散 (=  $X^2$  の平均 -  $X$  の平均の二乗)

$$V[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 = ? \quad (55)$$

## 5 指数分布 #2

指数分布に関しては、パラメータを  $\lambda$  とする記法の他に、次のように  $\mu = 1/\lambda$  をパラメータに用いる記法も用いられる。

確率密度関数

$$f(x) = \frac{d}{dx} \left( 1 - \exp\left(-\frac{x}{\mu}\right) \right) = ? \quad (56)$$

累積分布関数

$$F(k) = Pr[X \leq k] = 1 - \exp\left(-\frac{x}{\mu}\right) \quad (57)$$

$X$  の期待値 (=  $X$  の平均)

$$E[X] = \int_0^{\infty} xf(x) dx = ? \quad (58)$$

$X^2$  の期待値 (=  $X^2$  の平均)

$$E[X^2] = \int_0^{\infty} x^2 f(x) dx = ? \quad (59)$$

$X$  の分散 (=  $X^2$  の平均 -  $X$  の平均の二乗)

$$V[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 = ? \quad (60)$$

## 6 レポート略解

#4-3 参考書の略解は、次の計算を略記しているだけ。定義から

$$I = \int_{-\infty}^{\infty} (x - c) f_X(x) dx \quad (61)$$

$x = c - y$  と変数変換すると、 $y: \infty \rightarrow -\infty$  かつ  $dx = -dy$  となり

$$\int_{-\infty}^{\infty} (x - c) f_X(x) dx = \int_{\infty}^{-\infty} (-y) f_X(c - y) (-dy) = \int_{-\infty}^{\infty} y f_X(c - y) dy = - \int_{-\infty}^{\infty} y f_X(c - y) dy = - \int_{-\infty}^{\infty} y f_X(c + y) dy \quad (62)$$

更に  $y = z - c$  と変数変換すると、 $z: -\infty \rightarrow \infty$  かつ  $dy = dz$  となり

$$- \int_{-\infty}^{\infty} y f_X(c + y) dy = - \int_{-\infty}^{\infty} (z - c) f_X(z) dz = - \int_{-\infty}^{\infty} (x - c) f_X(x) dx = -I \quad (63)$$

を得る。これら二つの式より、 $I = -I$  となるから、 $I = 0$  でなければならない。よって

$$\int_{-\infty}^{\infty} (x - c) f_X(x) dx = 0 \Leftrightarrow \int_{-\infty}^{\infty} x f_X(x) dx = E[X] = c \quad (64)$$

## 7 レポート課題

確率分布  $F$  のモーメント母関数は、

$$E[e^{tX}] = \sum_{k \in \mathcal{X}} \exp(tk) p(k) \quad (65)$$

あるいは

$$E[e^{tX}] = \int_{x \in \mathcal{X}} \exp(tk) f(k) dx \quad (66)$$

で定義される。

#6-1  $n = 1, 2, 3$  の二項分布、それぞれについて、モーメント母関数を求めよ。

#6-2 上で求めた二項分布のモーメント母関数それぞれを、 $t = 0$  の回りでマクローリン展開せよ。ただし四次の項までで良い。

#6-3 指数分布のモーメント母関数を求めよ。

#6-4 上で求めた指数分布のモーメント母関数を、 $t = 0$  の回りでマクローリン展開せよ。ただし四次の項までで良い。

## 8 確率分布の計算に関するメモ

確率論に出てくる計算の種類は、実は多くはない。積分計算も、式変形や変数変換を施して、既知の積分に帰着させることが通例だが、それでも幾つかの量や積分には、慣れておくのが望ましい。その中の幾つかを次に掲げておく。

### 8.1 二項分布

組み合わせの数

$${}_n C_k = \frac{n!}{k!(n-k)!} = \binom{n}{k} \quad (67)$$

二項定理

$$(a+b)^n = \sum_{k=0}^n {}_n C_k a^k b^{n-k} = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \quad (68)$$

これらを、二項分布では期待値や分散の計算に用いる。期待値では

$$\begin{aligned} E[X] &= \sum_{k=0}^n k {}_n C_k p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n n \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p \cdot p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!((n-1)-l)!} p^l (1-p)^{(n-1)-l} \\ &= np \{p + (1-p)\}^{n-1} \\ &= np \end{aligned} \quad (69)$$

分散では、

$$V[X] = E[(X-\mu)^2] = E[X^2] - (E[X])^2 \quad (70)$$

のどちらでもなく、少し巧妙に、

$$V[X] = E[X(X-1)] + E[X] - (E[X])^2 \quad (71)$$

を用いると、計算が簡単になる。 $k^2$  をかけて総和を求めるより、 $k(k-1)$  をかけた総和を求める方が、二項定理に馴染みやすい。

$$\begin{aligned} E[X(X-1)] &= \sum_{k=0}^n k(k-1) {}_n C_k p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=2}^n k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \sum_{k=2}^n n(n-1) \frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} p^2 \cdot p^{k-2} (1-p)^{(n-2)-(k-2)} \\
&= n(n-1) p^2 \sum_{l=0}^{n-2} \frac{(n-2)!}{l!((n-2)-l)!} p^l (1-p)^{(n-2)-l} \\
&= n(n-1) p^2 \{p + (1-p)\}^{n-2} \\
&= n(n-1) p^2
\end{aligned} \tag{72}$$

より、

$$V[X] = n(n-1)p^2 + np - (np)^2 = np(1-p) \tag{73}$$

## 8.2 指数分布

指数分布の期待値の計算は、指数関数の積分

$$\int e^a x dx = e^{ax}/a + C \tag{74}$$

と、部分積分

$$\int_{x \in \mathcal{X}} g(x) e^x dx = [e^x g(x)]_{x \in \mathcal{X}} - \int_{x \in \mathcal{X}} g'(x) e^x dx \tag{75}$$

の組み合わせ、あるいは繰り返して乗り切れる。例えば  $X$  の期待値の計算は、

$$\begin{aligned}
E[X] &= \int_0^{\infty} x \lambda e^{-\lambda x} dx \\
&= [x(-e^{-\lambda x})]_{x=0}^{\infty} - \int_0^{\infty} (-e^{-\lambda x}) dx \\
&= 0 - 0 + \int_0^{\infty} e^{-\lambda x} dx \\
&= \left[ -\frac{1}{\lambda} e^{-\lambda x} \right]_{x=0}^{\infty} \\
&= 0 - \left( -\frac{1}{\lambda} \right) = \frac{1}{\lambda}
\end{aligned} \tag{76}$$

となる。ここで  $x \frac{1}{\lambda} e^{-\lambda x}$  の  $x \rightarrow \infty$  の極限は、 $x \rightarrow \infty$  よりも速く  $e^{-\lambda x} \rightarrow 0$  となることから 0 になるのは、既知とする。

2乗の期待値は、途中で  $X$  の期待値の  $2/\lambda$  倍になったことに気づけば、そこで計算はおしまい。

$$\begin{aligned}
E[X^2] &= \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx \\
&= [-x^2(-e^{-\lambda x})]_{x=0}^{\infty} - \int_0^{\infty} (-2xe^{-\lambda x}) dx \\
&= \int_0^{\infty} 2xe^{-\lambda x} dx \\
&= \frac{2}{\lambda} \int_0^{\infty} \lambda x e^{-\lambda x} dx \\
&= \frac{2}{\lambda} E[X] = \frac{2}{\lambda^2}
\end{aligned} \tag{77}$$

気づかなければ、3行目以降、部分積分をもう一度繰り返す。

$$\begin{aligned}
 E[X^2] &= \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx \\
 &= \left[ -x^2 (-e^{-\lambda x}) \right]_{x=0}^{\infty} - \int_0^{\infty} (-2xe^{-\lambda x}) dx \\
 &= 0 - 0 + \int_0^{\infty} 2xe^{-\lambda x} dx \\
 &= \left[ 2x \frac{-1}{\lambda} e^{-\lambda x} \right]_{x=0}^{\infty} - \int_0^{\infty} 2 \frac{-1}{\lambda} e^{-\lambda x} dx \\
 &= 0 - 0 + \int_0^{\infty} 2 \frac{1}{\lambda} e^{-\lambda x} dx \\
 &= \left[ \frac{2}{\lambda} \frac{-1}{\lambda} e^{-\lambda x} \right]_{x=0}^{\infty} \\
 &= 0 - \frac{-2}{\lambda^2} = \frac{2}{\lambda^2}
 \end{aligned} \tag{78}$$

いずれにせよ、分散は次のように求まる。

$$V[X] = E[X^2] - (E[X])^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2} \tag{79}$$

### 8.3 ガンマ分布

ガンマ関数

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx, \quad s > 0 \tag{80}$$

### 8.4 ベータ分布

ベータ関数

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx, \quad \alpha, \beta > 0 \tag{81}$$

### 8.5 マクローリン展開とテイラー展開

無限回微分可能な関数  $g(x)$  の  $k$  階の導関数を  $g^{(k)}$  とおく。まず、テイラーの公式

$$g(x) = \sum_{k=0}^{n-1} \frac{g^{(k)}(a)}{k!} (x-a)^k + R_n(x) \tag{82}$$

を既知とする。 $R_n$  はこの式の等号が成り立つように定められる剰余項である。ロルの定理から  $x$  に依存して決まり、区間  $(a, b)$  の値を取る関数  $c(x)$  を用いて

$$R_n(x) = \frac{g^{(n)}(c(x))}{n!} (x-a)^n \tag{83}$$

と表せる。

テイラーの公式で  $n \rightarrow \infty$  とせず、有限の  $n$  のまま、 $x = a$  の近傍で

$$g(x) \simeq \sum_{k=0}^n \frac{g^{(k)}(a)}{k!} (x-a)^k \tag{84}$$

とすると、 $g(x)$  の  $x = a$  の近傍での多項式近似が得られる。ただし、この近似が良い精度を持つためには、剰余項が

$$R_n(x) = o((x-a)^{n-1}) \quad (85)$$

を満たさなければならない。

$g(x)$  のマクローリン展開は、テイラーの公式で  $a=0$  と置き、さらに  $n \rightarrow \infty$  の極限

$$\begin{aligned} g(x) &= \sum_{k=0}^{\infty} \frac{g^{(k)}(0)}{k!} x^k \\ &= g(0) + \sum_{k=1}^{\infty} \frac{g^{(k)}(0)}{k!} x^k \\ &= g(0) + g'(0)x + \frac{g''(0)}{2}x^2 + \frac{g'''(0)}{6}x^3 + \dots \end{aligned} \quad (86)$$

として与えられる。マクローリン展開の展開の中心を  $x=0$  ではなく  $x=a$  とすると、テイラー展開と言われる。

$$\begin{aligned} g(x) &= \sum_{k=0}^{\infty} \frac{g^{(k)}(a)}{k!} (x-a)^k \\ &= g(a) + \sum_{k=1}^{\infty} \frac{g^{(k)}(a)}{k!} (x-a)^k \\ &= g(a) + g'(a)(x-a) + \frac{g''(a)}{2}(x-a)^2 + \frac{g'''(a)}{6}(x-a)^3 + \dots \end{aligned} \quad (87)$$

テイラー展開やマクローリン展開は、テイラーの公式の極限であり、これら存在する条件は、テイラーの公式の剰余項が

$$\lim_{n \rightarrow \infty} R_n(x) = 0 \quad (88)$$

を満たすことである。

## 9 レポート提出要領

下記の要領でレポートを作成し、提出すること。

課題番号	#6 (2012.05.24 出題)
提出期限	2012年5月28日 午後4時30分
提出場所	西5号館3階総合情報学科事務室前の集合ポストの「確率論」とある投函口
様式	A4もしくはB5 (ルーズリーフ可、両面可)
その他	丸写しは採点して飽きるし、剽窃は自分のためにならない 各自が自力で取り組むことを、切に願う 成書を参考にするなどは言わないが、参考にした書籍があれば、著者への礼儀として必ず記すこと 表紙はつけないこと 1ページ目の上部に、「講義名」「レポート番号」「学籍番号」「氏名」「投函日」を記すこと

下記は見本である。



提出日:2012/05/28

確率論レポート #6

学籍番号:0000000

氏 名:電通 大

課題#6-1

## 参考文献

- [1] 微分積分学の教科書.
- [2] 永田靖 (2005) 「統計学のための数学入門 30 講」朝倉書店.
- [3] 宮川雅巳 (1988) 「統計技法」共立出版.
- [4] 稲垣宣生 (2003) 「数理統計学」改訂版, 裳華房.