24 mai 2012
PROB #6

# 確率論 (*Probability Theory*) 第 6 週

w.yamamoto

## 1　累積分布関数

標本空間を $\mathcal{X}$ と書く。$\mathcal{X}$ 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。$\mathcal{X}$ 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。

離散確率分布については、標本空間 $\mathcal{X}$ の要素はしばしば整数で表されるので、その要素を $k$ と記すことにする。離散確率分布であれば、$\{k\}$ for $\forall k \in \mathcal{X}$ は標本空間のひとつの分割を与えるので、 $\forall k \in \mathcal{X}$ について確率

$$0 \le Pr\left[X = k\right] \le 1 \tag{1}$$

が定義できる。これを確率関数と呼び、

$$p\left(k\right) = Pr\left[X = k\right] \tag{2}$$

と書く。

連続確率分布については、標本空間 $\mathcal{X}$ の要素はしばしば実数で表されるので、その要素を $x$ と記すことにする。連続確率分布であれば、任意の単調実数列 $\inf \mathcal{X} a_0 < a_1 < a_2 < \cdots < a_n < \cdots$ は標本空間のひとつの分割

$$p\left(a_0, a_1\right], \left(a_1, a_2\right], \left(a_2, a_3\right], \ldots, \left(a_{n-1}, a_n\right], \left(a_n, a_{n+1}\right], \cdots \tag{3}$$

を与えるので、 $\forall k \in \mathcal{X}$ について確率

$$0 \le Pr\left[X \in \left(a_k, a_{k+1}\right]\right] \le 1 \tag{4}$$

が定義できる。ただしこれだと不便なので、上の部分集合を左から順に和集合を作った

$$p\left(a_0, a_1\right], \left(a_0, a_2\right], \left(a_0, a_3\right], \ldots, \left(a_0, a_n\right], \left(a_0, a_{n+1}\right], \cdots \tag{5}$$

と、互いに疎ではないが、

$$p\left(a_0, a_1\right] \subset \left(a_0, a_2\right] \subset \left(a_0, a_3\right] \subset \ldots \subset \left(a_0, a_n\right] \subset \left(a_0, a_{n+1}\right] \subset \cdots \tag{6}$$

という単調増加な集合の列に対して、確率を定める。

$$F\left(x\right) = Pr\left[X \le x\right] = Pr\left[X \in \left(-\infty, x\right]\right] \tag{7}$$

これを累積分布関数と言う。

累積分布関数は、離散確率分布についても同様に

$$F\left(x\right) = Pr\left[X \le x\right] = \sum_{k \le x, k \in \mathcal{X}} p\left(k\right) \tag{8}$$

と定めることはできる。

確率分布の累積分布関数が $F\left(x\right)$ あるいは $F\left(k\right)$ 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

$$Pr\left[X \in (a, b]\right] = Pr\left[X \in (-\infty, b]\right] - Pr\left[X \in (-\infty, a]\right] = F\left(b\right) - F\left(a\right) \tag{9}$$

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

$$f(x) = \frac{d}{dx} F(x) \tag{10}$$

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 $(x - \delta/2, x + \delta]$ に含まれる確率が

$$Pr[X \in (x - \delta x/2, x + \delta x/2]] \simeq f(x)\,\delta x \tag{11}$$

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。

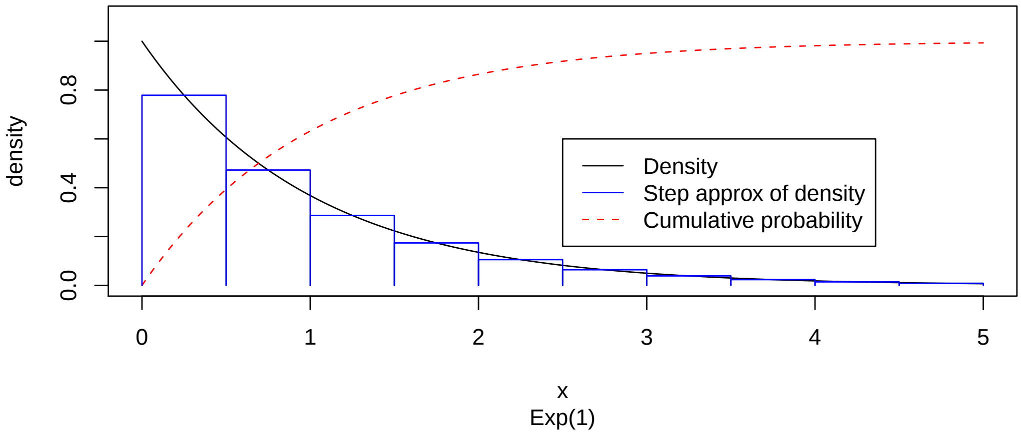

図 1 $\lambda = 1$ の指数分布の確率密度関数とその階段近似

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

$$F\left(x_1, x_2, \ldots, x_p\right) = Pr\left[X_1 \leq x_1\ and X_2 \leq x_2\ \ldots\ X_p \leq x_p\right] \tag{12}$$

またその確率密度関数は

$$f\left(x_1, x_2, \ldots, x_p\right) = \frac{d^p}{dx_1 dx_2 \cdots dx_p} F\left(x_1, x_2, \ldots, x_p\right) \tag{13}$$

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

$$E[X] = \begin{cases} \displaystyle\sum_{k \in \mathcal{X}} kp(k) & F \text{ が離散分布の場合} \\ \displaystyle\int_{x \in \mathcal{X}} xf(x)\,dx & F \text{ が連続分布の場合} \end{cases} \tag{14}$$

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

$$E[g(X)] = \begin{cases} \displaystyle\sum_{k \in \mathcal{X}} g(k)\,p(k) & F \text{ が離散分布の場合} \\ \displaystyle\int_{x \in \mathcal{X}} g(x)\,f(x)\,dx & F \text{ が連続分布の場合} \end{cases} \tag{15}$$

と定める。

- $E[X]$ : (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。しばしば $\mu$、$\mu_1$、あるいは $m_1$ と記される。
- $E\left[X^2\right]$ : (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、が平均して幾つぐらいの値をとるか。たまに $m_2$ と記される。
- $V[X] = E\left[(X - E[X])^2\right]$ : (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[1]でどのぐらい離れているか。しばしば $\sigma^2$ あるいは $\mu_2$ と表される。

$X^k$ の期待値 $E\left[X^k\right]$ 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、

$$m_k = E\left[X^k\right] \tag{16}$$

と記される。一方、$(X-\mu)^k$ の期待値 $E\left[(X-\mu)^k\right]$ 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、

$$\mu_k = E\left[(X-\mu)^k\right] \tag{17}$$

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。

$$\begin{aligned}
E[aX+b] &= \int_{x\in\mathcal{X}} (ax+b)\, f(x)\, dx \\
&= \int_{x\in\mathcal{X}} axf(x)\, dx + \int_{x\in\mathcal{X}} bf(x)\, dx \\
&= a\int_{x\in\mathcal{X}} xf(x)\, dx + b\int_{x\in\mathcal{X}} f(x)\, dx \\
&= aE[X] + b \\
E[aX+b] &= \sum_{k\in\mathcal{X}} (ak+b)\, p(k) \\
&= \sum_{k\in\mathcal{X}} akp(k) + \sum_{k\in\mathcal{X}} bp(k) \\
&= a\sum_{k\in\mathcal{X}} kp(x) + b\sum_{k\in\mathcal{X}} p(k) \\
&= aE[X] + b
\end{aligned} \tag{18}$$

この関係は例えば、

$$E\left[aX^2+b\right] = aE\left[X^2\right] + b \tag{19}$$

などと、$X$ の関数についても成り立つ。

$$E[ag(X)+b] = aE[g(X)] + b \tag{20}$$

分散 ($X-\mu$ の二乗の期待値) についても、

$$\begin{aligned}
V[X] &= \sigma^2 \\
&= E\left[(X-\mu)^2\right] = \int_{x\in\mathcal{X}} (x-\mu)^2 f(x)\, dx
\end{aligned}$$

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

$$
\begin{aligned}
&= \int_{x\in\mathcal{X}} x^2 f(x)\,dx - 2\int_{x\in\mathcal{X}} x\mu f(x)\,dx + \int_{x\in\mathcal{X}} \mu^2 f(x)\,dx \\
&= E\left[X^2\right] - 2\mu \int_{x\in\mathcal{X}} x f(x)\,dx + \mu^2 \int_{x\in\mathcal{X}} f(x)\,dx \\
&= E\left[X^2\right] - 2\mu E[X] + \mu^2 \\
&= E\left[X^2\right] - (E[X])^2\, m_2 - {m_1}^2 \qquad (21)
\end{aligned}
$$

$$
\begin{aligned}
V[X] &= \sigma^2 \\
&= E\left[(X-\mu)^2\right] = \sum_{x\in\mathcal{X}} (x-\mu)^2\, p(x) \\
&= \sum_{x\in\mathcal{X}} x^2 p(x) - 2\sum_{x\in\mathcal{X}} x\mu p(x) + \sum_{x\in\mathcal{X}} \mu^2 p(x) \\
&= E\left[X^2\right] - 2\mu \sum_{x\in\mathcal{X}} x p(x) + \mu^2 \sum_{x\in\mathcal{X}} p(x) \\
&= E\left[X^2\right] - 2\mu E[X] + \mu^2 \\
&= E\left[X^2\right] - (E[X])^2\, m_2 - {m_1}^2 \qquad (22)
\end{aligned}
$$

および、任意の定数 $a$ について

$$
\begin{aligned}
V[aX] &= E\left[(aX - a\mu)^2\right] = \int_{x\in\mathcal{X}} (ax - a\mu)^2 f(x)\,dx \\
&= \int_{x\in\mathcal{X}} a^2 (x-\mu)^2 f(x)\,dx \\
&= a^2 \int_{x\in\mathcal{X}} (x-\mu)^2 f(x)\,dx \\
&= a^2 V[X] = a^2\sigma^2 \qquad (23)
\end{aligned}
$$

$$
\begin{aligned}
V[aX] &= E\left[(aX - a\mu)^2\right] = \sum_{x\in\mathcal{X}} (ax - a\mu)^2 f(x)\,dx \\
&= \sum_{x\in\mathcal{X}} a^2 (x-\mu)^2 f(x)\,dx \\
&= a^2 \sum_{x\in\mathcal{X}} (x-\mu)^2 f(x)\,dx \\
&= a^2 V[X] = a^2\sigma^2 \qquad (24)
\end{aligned}
$$

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

$$
p(0) = Pr[X=0] = 1-p,\ p(1) = Pr[X=1] = p \qquad (25)
$$

となる。毎回の試行を独立[2]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

$$
\begin{aligned}
p(0) &= Pr[X_1 = 0\ and\ X_2 = 0] \\
&= Pr[X_1 = 0]\, Pr[X_2 = 0] = (1-p)^2 \qquad (26)
\end{aligned}
$$

[2]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。$Pr[X_1 \le a\ and\ X_2 \le b] = Pr[X_1 \le a]\, Pr[X_2 \le b]$ も、$Pr[X_1 = a\ and\ X_2 = b] = Pr[X_1 \le a]\, Pr[X_2 \le b]$ も成り立つ。

$$
\begin{aligned}
p(1) &= Pr[\{X_1 = 0 \; and \; X_2 = 1\} \; or \; \{X_1 = 1 \; and \; X_2 = 0\}] \\
&= Pr[X_1 = 1] \, Pr[X_2 = 0] + Pr[X_1 = 0] \, Pr[X_2 = 1] \\
&= p(1-p) + (1-p)p = 2p(1-p) \\
&= {}_2C_1 p(1-p) \qquad (27) \\
p(2) &= Pr[X_1 = 1 \; and \; X_2 = 1] \\
&= Pr[X_1 = 1] \, Pr[X_2 = 1] = p^2 \qquad (28)
\end{aligned}
$$

$n$ 回の互いに独立な試行で $k$ 回の成功を観測する確率は、

$$
p(k) = Pr[X = k] = {}_nC_k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \qquad (29)
$$

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。

$$
\begin{aligned}
(p + (1-p))^n &= 1 \\
&= {}_nC_0 p^0 (1-p)^{n-0} + {}_nC_1 p^1 (1-p)^{n-1} + {}_nC_2 p^2 (1-p)^{n-2} + \cdots \\
&\quad + {}_nC_{n-1} p^{n-1} (1-p)^1 + {}_nC_n p^n (1-p)^0 \qquad (30)
\end{aligned}
$$

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

$$
p(k) = p(k; n, p) = Pr[X = k] = {}_nC_k p^k (1-p)^{n-k} \qquad (31)
$$

累積分布関数

$$
F(k) = Pr[X \le k] = \sum_{l=0}^{k} {}_nC_l p^l (1-p)^{n-l} \qquad (32)
$$

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。

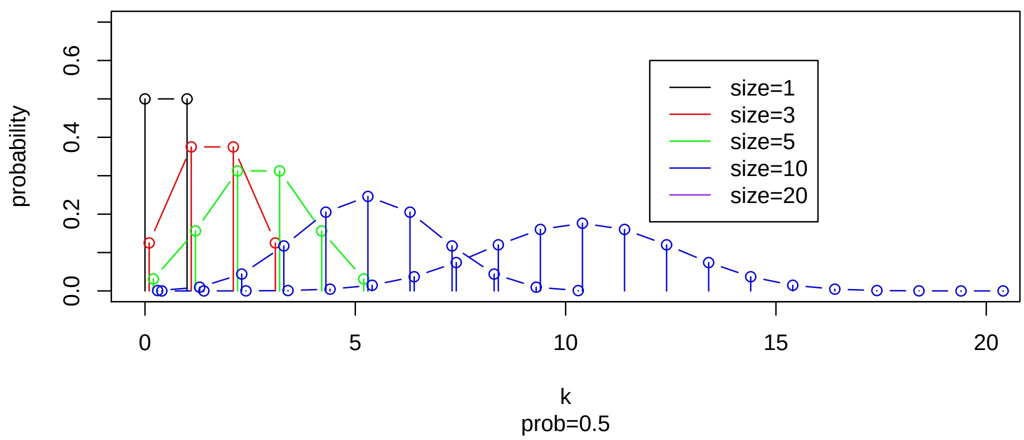


図 2 $p = 0.5$ の二項分布の確率関数 ($n = 1, 3, 5, 10, 20$)

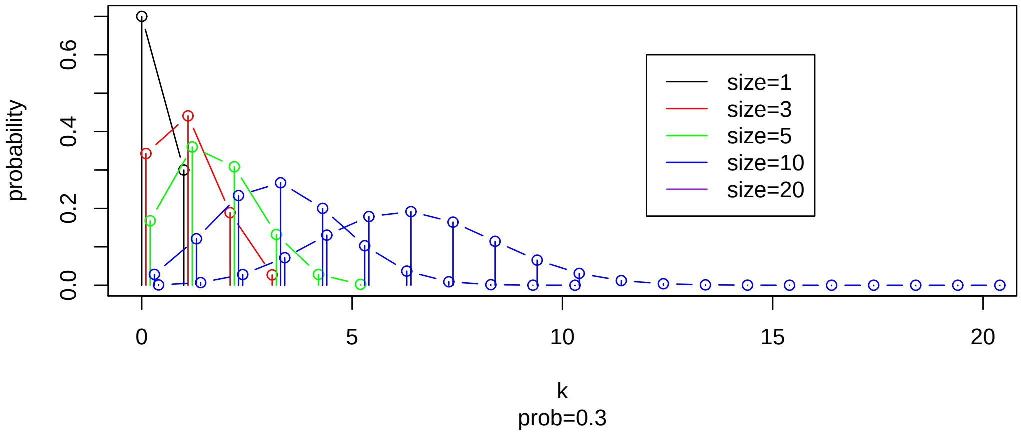


図 3 $p = 0.3$ の二項分布の確率関数 ($n = 1, 3, 5, 10, 20$)

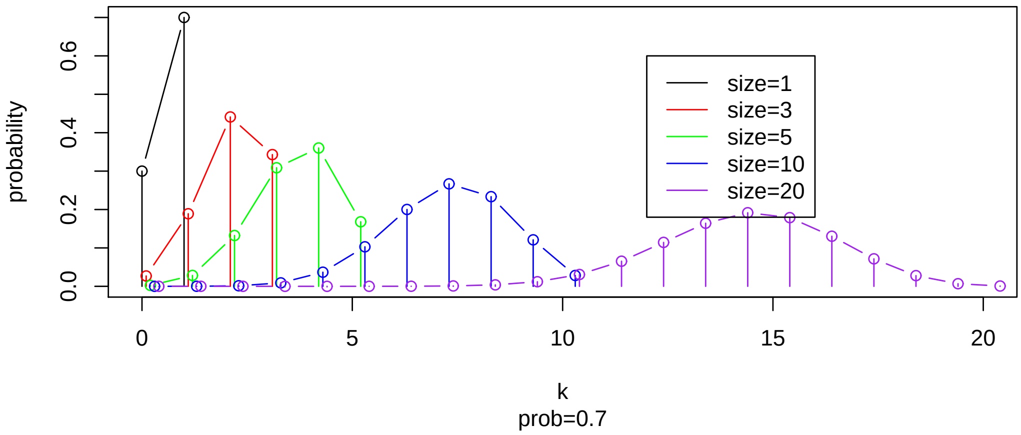


図 4 $p = 0.7$ の二項分布の確率関数 ($n = 1, 3, 5, 10, 20$)

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、

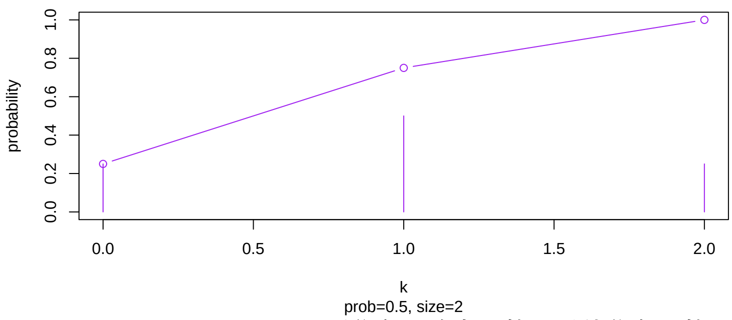


図 5 $n = 2, p = 0.5$ の二項分布の確率関数と累積分布関数

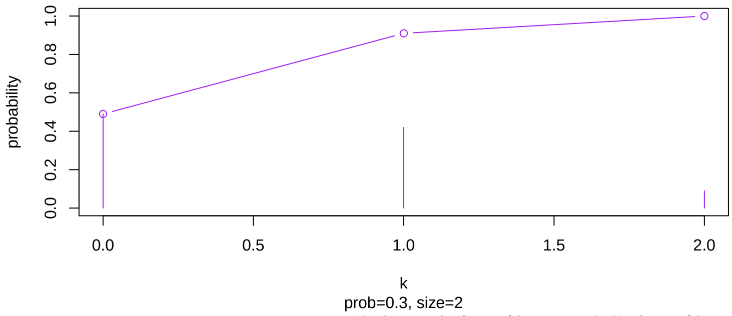

図 6 $n = 2, p = 0.3$ の二項分布の確率関数と累積分布関数

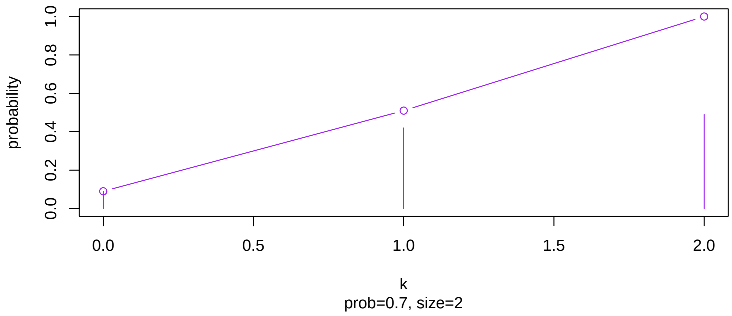

図 7 $n = 2, p = 0.7$ の二項分布の確率関数と累積分布関数

のように、自分で計算して描くことも簡単であろう。総回数が 20 の場合は、

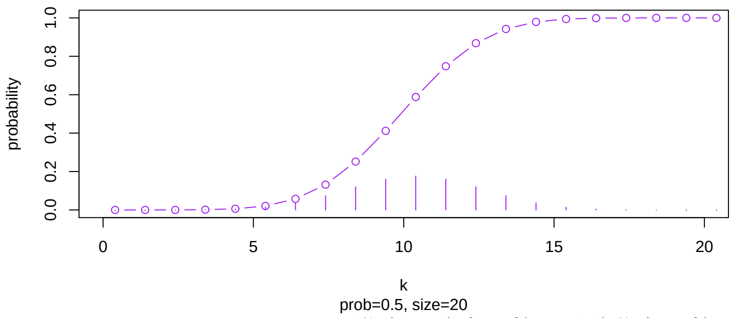

図 8 $n = 20, p = 0.5$ の二項分布の確率関数と累積分布関数

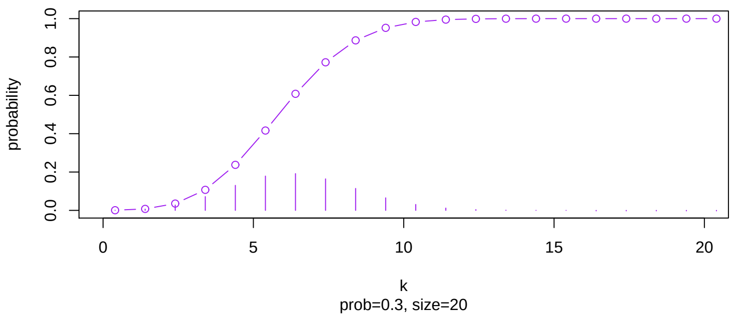

図 9 $n = 20$, $p = 0.3$ の二項分布の確率関数と累積分布関数

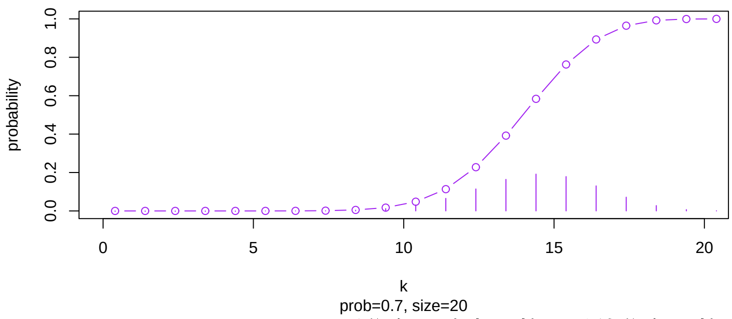


図 10 $n = 20$, $p = 0.7$ の二項分布の確率関数と累積分布関数

となる。

$X$ の期待値 (= $X$ の平均)

$$E[X] = \sum_{k=0}^{n} k_n C_k p^k (1-p)^{n-k} = \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = ? \tag{33}$$

$X^2$ の期待値 (= $X^2$ の平均)

$$E\left[X^2\right] = \sum_{k=0}^{n} k^2{}_n C_k p^k (1-p)^{n-k} = \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = ? \tag{34}$$

$X$ の分散 (= $X^2$ の平均 - $X$ の平均の二乗)

$$V[X] = E\left[(X - E[X])^2\right] = E\left[X^2\right] - (E[X])^2 = ? \tag{35}$$

## 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

$$f(x) = \frac{d}{dx}(1 - \exp(-\lambda x)) = ? \tag{36}$$

累積分布関数

$$F(k) = Pr[X \leq k] = 1 - \exp(-\lambda x) \tag{37}$$

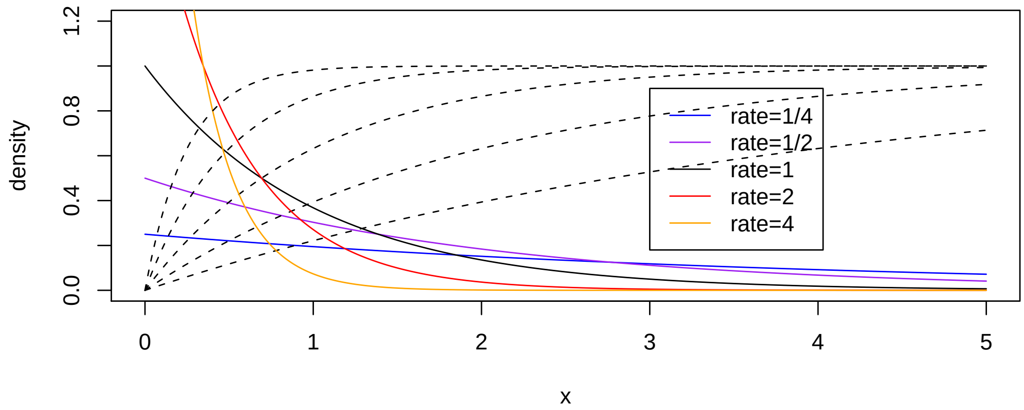


図 11 指数分布の累積分布関数と確率密度関数

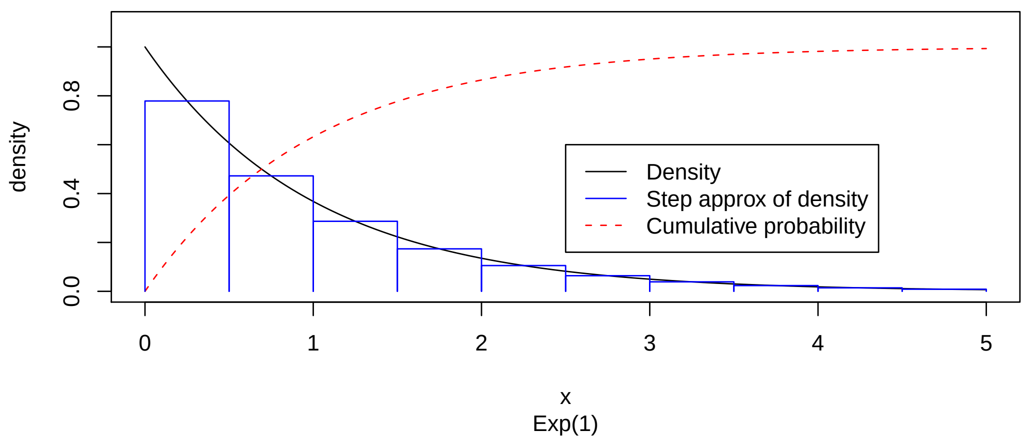


図 12 $\lambda = 1$ の指数分布の累積分布関数と確率密度関数

$X$ の期待値 (= $X$ の平均)

$$E[X] = \int_0^\infty x f(x) = ? \tag{38}$$

$X^2$ の期待値 (= $X^2$ の平均)

$$E[X] = \int_0^\infty x^2 f(x) = ? \tag{39}$$

$X$ の分散 (= $X^2$ の平均 - $X$ の平均の二乗)

$$V[X] = E\left[(X - E[X])^2\right] = E\left[X^2\right] - (E[X])^2 = ? \tag{40}$$

## 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

$$f(x) = \frac{d}{dx}\left(1 - \exp\left(-\frac{x}{\mu}\right)\right) = ? \tag{41}$$

累積分布関数

$$F(k) = Pr[X \le k] = 1 - \exp\left(-\frac{x}{\mu}\right) \tag{42}$$

$X$ の期待値 (= $X$ の平均)

$$E[X] = \int_0^\infty x f(x) = ? \tag{43}$$

$X^2$ の期待値 (= $X^2$ の平均)

$$E[X] = \int_0^\infty x^2 f(x) = ? \tag{44}$$

$X$ の分散 (= $X^2$ の平均 - $X$ の平均の二乗)

$$V[X] = E\left[(X - E[X])^2\right] = E\left[X^2\right] - (E[X])^2 = ? \tag{45}$$

# 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

$$I = \int_{-\infty}^{\infty} (x - c) f_X(x)\, dx \tag{46}$$

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

$$\int_{-\infty}^{\infty} (x - c) f_X(x)\, dx = \int_{\infty}^{-\infty} (-y) f_X(c - y)(-dy) = \int_{\infty}^{-\infty} (y) f_X(c - y)\, dy = -\int_{-\infty}^{\infty} (y) f_X(c - y)\, dy = -\int_{-\infty}^{\infty} (y) f_X(c + y)\, dy \tag{47}$$

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

$$-\int_{-\infty}^{\infty} (y) f_X(c + y)\, dy = -\int_{-\infty}^{\infty} (z - c) f_X(z)\, dz = -\int_{-\infty}^{\infty} (x - c) f_X(x)\, dx = -I \tag{48}$$

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

$$\int_{-\infty}^{\infty} (x - c) f_X(x)\, dx = 0 \iff \int_{-\infty}^{\infty} x f_X(x)\, dx = E[X] = c \tag{49}$$

# 7 レポート課題

確率分布 $F$ のモーメント母関数は、

$$E\left[e^{tX}\right] = \sum_{k \in \mathcal{X}} \exp(tk)\, p(k) \tag{50}$$

あるいは

$$E\left[e^{tX}\right] = \int_{x \in \mathcal{X}} \exp(tk)\, f(k)\, dx \tag{51}$$

で定義される。

**#6-1** 二項分布の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。

組み合わせの数

$$ {}_{n}C_{k} = \frac{n!}{k!\,(n-k)!} \tag{52} $$

指数関数の積分

$$ \int e^{a}xdx = e^{ax}/a + C \tag{53} $$

ガンマ関数

$$ \Gamma(s) = \int_{0}^{\infty} x^{s-1}e^{-x}dx,\ s>0 \tag{54} $$

ベータ関数

$$ B(\alpha,\beta) = \int_{0}^{1} x^{\alpha-1}(1-x)^{\beta-1}\,dx,\ \alpha,\beta>0 \tag{55} $$

## 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 | #4 (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 | 丸写しは採点していて飽きるし、剽窃は自分のためにならない<br>各自が自力で取り組むことを、切に願う<br>成書を参考にするなとは言わないが、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<br>表紙はつけないこと<br>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