大数の法則・正規分布・中心極限定理

樋口さぶろお

龍谷大学理工学部数理情報学科

確率統計☆演習 I　L09(2016-12-01 Thu)

最終更新: Time-stamp: "2016-12-16 Fri 19:36 JST hig"

今日の目標

- 大数の法則の意味が説明できる 塚田確率統計 §5.2
- 正規分布の母平均値・母分散・確率が積分や表で求められる. 塚田確率統計 §4.7
- 中心極限定理の意味が説明でき, 確率の計算に利用できる. 塚田確率統計 §5.3

http://hig3.net

L08-Q1

Quiz 解答:離散型確率変数の独立性

確率の和は 1 なので, $\frac{2}{12}+\frac{1}{12}+A+B=1$.
よって,

$$f_{\mathrm{Y}}(y)=\begin{cases}\frac{3}{12} & (y=3)\\ \frac{9}{12} & (y=7)\end{cases}$$

独立性から,

$$\begin{aligned}f_{\mathrm{XY}}(2,3)=&f_{\mathrm{X}}(2)\frac{3}{12}=\frac{2}{12},\\ f_{\mathrm{XY}}(3,3)=&f_{\mathrm{X}}(3)\frac{3}{12}=\frac{1}{12},\\ f_{\mathrm{XY}}(2,7)=&f_{\mathrm{X}}(2)\frac{9}{12}=A,\\ f_{\mathrm{XY}}(3,7)=&f_{\mathrm{X}}(3)\frac{9}{12}=B.\end{aligned}$$

$A, B, f_{\mathrm{X}}(2), f_{\mathrm{X}}(3)$ を未知数として解くと, $A=\frac{6}{12}$, $B=\frac{3}{12}$.

L08-Q2

Quiz 解答:独立な確率変数の期待値

1. X, Y は独立なので, $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$ であることに注意して, $\mathrm{E}[(-2X+3Y)(X+5Y)] = \mathrm{E}[-2X^2] + \mathrm{E}[-7XY] + \mathrm{E}[15Y^2] = -2(\mathrm{V}[X] + \mathrm{E}[X]^2) - 7\mathrm{E}[X]\mathrm{E}[Y] + 15(\mathrm{V}[Y] + \mathrm{E}[Y^2])$
2. 独立なので, $\mathrm{V}[XY] = \mathrm{V}[X] + \mathrm{V}[Y]$ であることに注意して, $\mathrm{V}[-2X+3Y] = \mathrm{V}[-2X] + \mathrm{V}[3Y] = 4\mathrm{V}[X] + 9\mathrm{V}[Y]$.

L08-Q4

Quiz 解答:二項分布

1. 二項分布 $\mathrm{B}(100, \frac{2}{3})$ に従う確率変数を X とする $P(X=50)$ を求めればいいから, ${}_{100}\mathrm{C}_{50}p^{50}(1-p)^{100-50} = \frac{100!}{50!50!}(\frac{2}{3})^{50}(1-\frac{2}{3})^{50}$.
2. $\mathrm{E}[X] = n \times p = \frac{200}{3}$.
3. $\mathrm{V}[X] = n \times p(1-p) = \frac{200}{9}$.

L08-Q5

Quiz 解答:ベルヌーイ分布

1. ベルヌーイ分布 $\mathrm{B}(1, 0.05)$ に従う確率変数を X とすると，
   $Y = 1000X$.
2. $\mathrm{E}[Y] = \mathrm{E}[1000X] = 1000\mathrm{E}[X] = 1000p = 50$.
   $\mathrm{V}[Y] = \mathrm{V}[1000X] = 1000^2\mathrm{V}[X] = 1000^2 p(1-p) = 47500$.

L09-Q6

Quiz 解答:独立同分布にしたがう変数の和　　　母平均値と母分散は，
$\mathrm{E}[aX + bY + c] = a\mathrm{E}[X] + b\mathrm{E}[Y] + c$, 独立分布の
$\mathrm{V}[aX + bY + c] = a^2\mathrm{V}[X] + b^2\mathrm{V}[Y]$ の式を繰り返し使えば求められる．

1. A は母平均値が 3, 母分散が $\frac{7}{100}$.
2. B は母平均値が 0, 母分散が 7.
3. C は母平均値が 0, 母分散が 1.

復習

チェビシェフの不等式 Chebyshev's inequality 塚田確率統計 3.5

X: 離散型または連続型確率変数
$\mu = \mathrm{E}[X]$: 母平均値
$\sigma^2 = \mathrm{V}[X]$: 母分散
$a > 0$: 任意の正の実数
のとき次が成立する.

$$P(|X - \mu| \geq a\sigma) \leq \frac{1}{a^2}$$

要するに

独立同分布の性質 塚田確率統計 5.1

独立同分布 (i.i.d.)

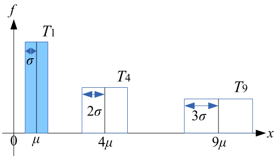
離散型/連続型確率変数 $X_1, X_2, \ldots, X_n$ が, たがいに独立で, すべて同じ確率分布に従う (同じ確率関数 $f(x)$) とする.
これを $X_1, \ldots, X_n$ は独立同分布に従う (i.i.d.=independent and identically-distributed) という.

新しい確率変数: $U_n = X_1 + \cdots + X_n$
母平均値 $\mathrm{E}[X_i] = \mu$, 母分散 $\mathrm{V}[X_i] = \sigma^2$ としたとき,

$$\mathrm{E}[U_n] = \sum_{i=1}^{n} \mathrm{E}[X_i] = n \times \mu.$$

$$\mathrm{V}[U_n] = \sum_{i=1}^{n} \mathrm{V}[X_i] = n \times \sigma^2.$$

U_n の確率密度関数はこんな感じ?

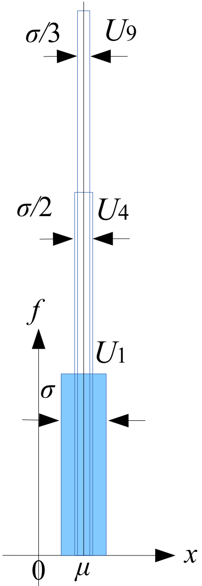


W_n の確率密度関数はこんな感じ?

新しい確率変数: $W_n = \frac{1}{n}U_n = \frac{1}{n}(X_1 + \cdots + X_n)$

$$\mathrm{E}[W_n] = \mathrm{E}\left[\tfrac{1}{n}U_n\right] = \frac{1}{n} \times n \times \mu.$$

$$\mathrm{V}[W_n] = \mathrm{V}\left[\tfrac{1}{n}U_n\right] = \left(\frac{1}{n}\right)^2 \times n \times \sigma^2.$$



大数の (弱) 法則 塚田確率統計 5.2

$W_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ は $\forall \epsilon > 0 \ \lim_{n\to+\infty} P(|W_n - \mu_{\mathrm{X}}| \geq \epsilon) = 0$ を満たす.

つまり, n 大で W_n は母平均値 $\mu_{\mathrm{X}} = \mathrm{E}[X]$ に「必ず近い」(W_n が μ_X に確率収束)

証明 $\mu_{W_n} = \mu_X, \sigma^2_{W_n} = \sigma^2_X/n$. W_n に対するチェビシェフの不等式より,

$$P(|W_n - \mu_{\mathrm{X}}| \geq a \times \frac{\sigma_{\mathrm{X}}}{\sqrt{n}}) \leq \frac{1}{a^2}$$

$a = \frac{\epsilon}{\frac{\sigma_{\mathrm{X}}}{\sqrt{n}}}$ とすると, $n \to +\infty$ で

$$P(|W_n - \mu_{\mathrm{X}}| \geq \epsilon) \leq \frac{\sigma^2_{\mathrm{X}}}{\epsilon^2 n} \to 0$$

これが母期待値の直観的意味. 要するに,

ここまで来たよ

1. 独立性・二項分布・ベルヌーイ分布
2. 大数の法則・正規分布・中心極限定理
   - 正規分布
   - 中心極限定理

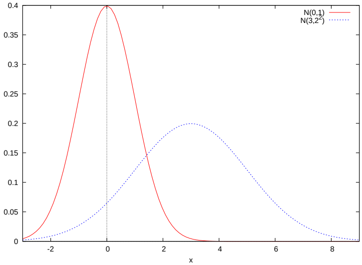
一般の正規分布 塚田確率統計 §4.7 高校 数学 B

正規＝normal

(一般の) 正規分布 $\mathrm{N}(b, a^2)$ の確率密度関数

$$f(x; b, a^2) = \frac{1}{\sqrt{2\pi a^2}} \mathrm{e}^{-\frac{(x-b)^2}{2a^2}}.$$

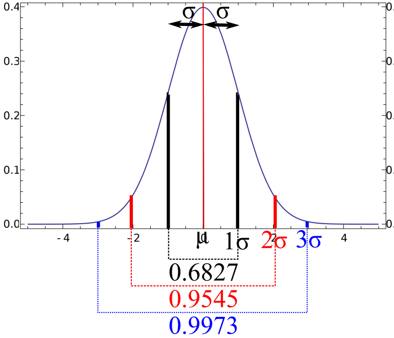
b, a^2:パラメタ

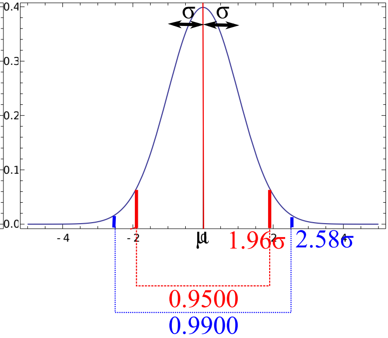


難しいので，まず $b = 0, a = 1$ の場合を考える．
$z = \frac{x-b}{a}$ または $x = az + b$
$y = f(z; 0, 1)$ のグラフを，横に a 倍，横に b だけ平行移動して，縦に $1/\sqrt{a^2}$ 倍したものが $y = f(x; b, a^2)$

標準正規分布 $\mathrm{N}(0, 1^2)$ の確率密度関数

$$f(z; 0, 1^2) = f(z) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z^2}{2}}.$$





標準正規分布 $\mathrm{N}(0,1^2)$ の性質

$P(c<Z<d)=$簡単じゃない．表 塚田確率統計付録 B, p207 をひけ．

$\mathrm{E}[1]=1,$ 塚田確率統計 p.83 微積分 II

$\mathrm{E}[Z]=0,$ 奇関数．塚田確率統計 p.85

$\mathrm{V}[Z]=1$ 塚田確率統計 p.85 微積分 II

$I(d)=\int_0^d f(z;0,1^2)\,\mathrm{d}z=P(0<Z<d)$ の表 塚田確率統計付録 B, p207 高校 数学 B

この $I(z)$ の表，$I(+\infty)=\frac{1}{2}$，Z の確率密度関数 f が偶関数であること，ですべての区間の確率が求められる．

上側確率 $Q(z)=\frac{1}{2}-I(z)$ を載せている本も多い．

L09-Q1

Quiz(標準正規分布の確率)

Z は標準正規分布 $\mathrm{N}(0, 1^2)$ に従う. $Z < -2$ となる確率は?

L09-Q2

Quiz(標準正規分布の確率)

Z は標準正規分布 $\mathrm{N}(0, 1^2)$ に従う連続型確率変数である.

1. 母期待値 $\mathrm{E}[Z^2]$ を求めよう.
2. 確率 $P(-0.56 < Z < +1.23)$ を数表から求めよう.

L09-Q3 $\boxed{\text{塚田確率統計 §4.13 問 3}}$

正規分布ふたたび

$Z = \frac{X-b}{a}$ または $X = aZ + b$, $Z \sim \mathrm{N}(0, 1^2)$.

$$
\begin{aligned}
\mathrm{E}[1] =& 1, \\
\mu_{\mathrm{X}} = \mathrm{E}[X] =& \mathrm{E}[aZ+b] = b, \\
\sigma_{\mathrm{X}}^2 = \mathrm{V}[X] =& \mathrm{V}[aZ+b] = a^2,
\end{aligned}
$$

塚田確率統計 p.83 微積分 II
塚田確率統計 p.85
塚田確率統計 p.85 微積分 II

つまり, $b = \mu_{\mathrm{X}}, a^2 = \sigma_{\mathrm{X}}^2$ ってこと.

(一般の) 正規分布 $\mathrm{N}(\mu, \sigma^2)$

母平均値 μ, 母分散 σ^2 の正規分布 $\mathrm{N}(\mu, \sigma^2)$ の確率密度関数は,

$$
f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.
$$

L09-Q4

Quiz(正規分布の確率)

連続型確率変数 X は, 確率密度関数

$$f(x) = \frac{1}{\sqrt{2\pi 3^2}} \mathrm{e}^{-\frac{(x-4)^2}{2\cdot 3^2}}$$

にしたがう.

1. $\mathrm{E}[X]$ を求めよう.
2. $\mathrm{V}[X]$ を求めよう.

L09-Q5

Quiz(正規分布の確率)

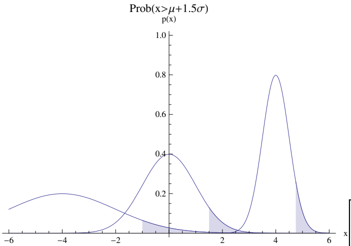
連続型確率変数 X は, 確率密度関数

$$f(x) = C \cdot \mathrm{e}^{-\frac{1}{18}x^2 + \frac{4}{9}x}$$

を持つ.

1. $\mathrm{E}[X]$ を求めよう.
2. $\mathrm{V}[X]$ を求めよう.
3. 定数 C を求めよう.

$\mathrm{N}(\mu, \sigma^2)$ の確率の求め方 I



対応する $Z = \frac{X-\mu_X}{\sigma_X} \sim \mathrm{N}(0, 1^2)$ の範囲を考えて, 表から求める.

一般の正規分布の確率

$X \sim \mathrm{N}(\mu_X, \sigma_X^2), \quad Z \sim \mathrm{N}(0, 1^2)$ のとき

$$P(c < X < d) = P\left(\frac{c-\mu_X}{\sigma} < \frac{X-\mu_X}{\sigma} < \frac{d-\mu_X}{\sigma}\right) = P\left(\frac{c-\mu_X}{\sigma} < Z < \frac{d-\mu_X}{\sigma}\right)$$

L09-Q6

Quiz(正規分布の確率)

X が母平均値 3, 母分散 4 の正規分布にしたがうとする.

1. $X \geq 5$ となる確率を求めよう.
2. $+1 \leq X \leq 7$ となる確率を求めよう.

L09-Q7

Quiz(正規分布の確率)

1. 母平均値 0, 母分散 1^2 の正規分布で, $0.5 \leq X < 0.7$ となる確率を求めよう.
2. 母平均値 0, 母分散 2^2 の正規分布で, $0.5 \leq X < 0.7$ となる確率を求めよう.
3. 母平均値 3, 母分散 2^2 の正規分布で, $4.0 \leq X < 4.4$ となる確率を求めよう.

L09-Q8 塚田確率統計 §4.13 問 4

ここまで来たよ

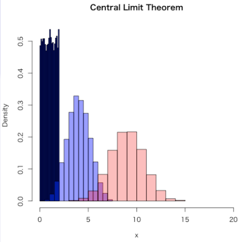
1. 独立性・二項分布・ベルヌーイ分布
2. 大数の法則・正規分布・中心極限定理
   - 正規分布
   - 中心極限定理

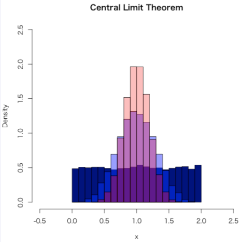
最初が一様分布でも，実は $n \to \infty$ で U や W の確率密度関数の形は長方形から崩れていく．分布の個性が消える! っていうか美しい形に!

中心極限定理 (いいかげんバージョン) 塚田確率統計 5.3

$X_1, \ldots, X_n$ が母平均値 μ, 母分散 σ^2 の独立同分布に従うとき,

- $U_n = X_1 + \cdots + X_n$, の確率分布は,
  の ☐ に似る
- $W_n = \frac{1}{n}(X_1 + \cdots + X_n)$ の確率分布は,
  ☐ に似る
- $Z_n = \frac{W_n - \mu}{\sigma\sqrt{n}}$ の確率分布は,
  ☐ に似る





中心極限定理 (厳密バージョン)

確率変数 $X_1, X_2, \ldots, X_n$ が, 母平均値 μ, 母分散 σ^2 の独立同分布に従うとする. 正規分布じゃない. どんな分布でも可

$Z_n = \frac{\frac{1}{n}(X_1+\cdots+X_n)-\mu}{\sigma} \times \sqrt{n}$ とすると,
Z_n は, $n \to +\infty$ の極限で, $\mathrm{N}(0, 1^2)$ に従う. すなわち

$$\lim_{n\to+\infty} P(a \leq Z_n < b) = \int_a^b \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x$$

「Z_n は $\mathrm{N}(0, 1^2)$ にしたがう Z に法則収束する」

法則収束とは, 関数列がある関数に収束すること.

証明

$\mathrm{E}[Z_n] = 0, \mathrm{V}[Z_n] = 1$ はすぐわかるが…

モーメント母関数を使うと瞬殺

確率統計☆演習 II

L09-Q9

Quiz(中心極限定理)

確率変数 $X_1, \ldots, X_{100}$ は $\mathrm{E}[X_i] = 1, \mathrm{V}[X_i] = \frac{1}{4}$ の独立同分布に従う. 次の確率変数の母平均値と母分散を求めよう. また, $n = 10$ が十分に大きいと思って, 指定の確率を求めよう.

1. 確率変数 $U = X_1 + X_2 + X_3 + \cdots + X_{100}$. 確率 $P(U > 110)$.
2. 確率変数 $W = \frac{1}{100}(X_1 + X_2 + X_3 + \cdots + X_{100})$. 確率 $P(W < 1.01)$.

連絡

- 予習問題は, 次々回の授業直前 を締切 (そこまでの最高点を記録) とします. でも, Trial までにやったほうが効率いいと思う.
  前からそうだけど, 予習問題が満点だと, Trial の満点の 1/3 まで保証されます.
- 配布資料は 1-503 向かいの引出, `http://hig3.net` で再配布.
- 加減乗除と平方根 (ルート) の使える電卓持ってきてね. 関数電卓でなくてもいいです. 携帯電話の機能・アプリでもかまいません.
- 樋口オフィスアワー木 6 金昼 (1-502), Math ラウンジ月-木昼 (1-614)
- 次回は 塚田確率統計 §6.1 塚田確率統計 §6.2 塚田確率統計 §7.1 塚田確率統計 §7.2
- Manaba にレポートあります. 2016-12-08 まで.

`https://manaba.ryukoku.ac.jp`