二項分布, 独立同分布の和

樋口さぶろお　https://hig3.net

龍谷大学理工学部数理情報学科

確率統計☆演習 I　L09(2019-12-02 Mon)

最終更新: Time-stamp: "2019-12-03 Tue 08:54 JST hig"

今日の目標

- 一般の正規分布の確率が積分や表で求められる 岩薩林 確率・統計 §4.5
- 二項分布の母ナントカが求められる 岩薩林 確率・統計 §3.4
- 独立同分布の和の母ナントカが求められる

L08-Q1
Quiz 解答:一様分布

1. $$f_{\mathrm{Z}}(z) = \frac{1}{2\sqrt{3}} \begin{cases} 1 & (-\sqrt{3} \le y < \sqrt{3}) \\ 0 & (\text{他}) \end{cases}$$
   より,
   $$f_{\mathrm{X}}(x) = \frac{1}{2\sqrt{3}} \begin{cases} 1 & (1 - 2\sqrt{3} \le y < 1 + 2\sqrt{3}) \\ 0 & (\text{他}) \end{cases}.$$
   すなわち, $X \sim \mathrm{U}(1 - 2\sqrt{3}, 1 + 2\sqrt{3})$.
2. $\mathrm{E}[X] = \frac{(1-2\sqrt{3})+(1+2\sqrt{3})}{2}$, または, $\mathrm{E}[X] = 2\mathrm{E}[Z] = 1$.
3. $\mathrm{V}[X] = \frac{((1+2\sqrt{3})-(1-2\sqrt{3}))^2}{12}$, または, $\mathrm{V}[X] = 2^2\mathrm{V}[Z] = 2^2$.

L08-Q2

Quiz 解答:標準正規分布の確率

確率密度関数が偶関数であることに注意する.

1. $P(-\infty < Z < 0) == \int_{-\infty}^{0} f(z;0,1^2)\,\mathrm{d}z = -I(-\infty) - I(0) = I(\infty) - 0 = \frac{1}{2}$.
   $\frac{1}{2} = 0.5$.
2. $P(-0.56 < Z < +1.23) = \int_{-0.56}^{1.23} f(z;0,1^2)\,\mathrm{d}z = -I(-0.56) + I(1.23) = I(0.56) + I(1.23)$ 表より,
   $0.2123 + 0.3907 = 0.6030$.

L08-Q3

Quiz 解答:正規分布の確率

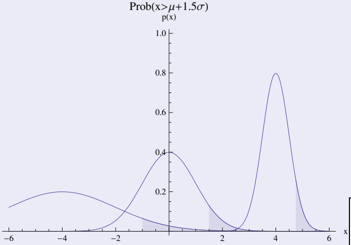
1. $x = 4$ を真ん中に幅 3 くらいの正規分布の確率密度関数のグラフ.
2. $\mathrm{E}[X] = 4$.
3. $\mathrm{V}[X] = 3^2$.

正規分布 $\mathrm{N}(\mu, \sigma^2)$ の確率の求め方 I

一般の正規分布の確率

$X \sim \mathrm{N}(\mu, \sigma^2)$ の確率は, 標準化 $Z = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0, 1^2)$ して求める.

$$P(c < X < d) = P(\tfrac{c-\mu}{\sigma} < \tfrac{X-\mu}{\sigma} < \tfrac{d-\mu}{\sigma}) = P(\tfrac{c-\mu}{\sigma} < Z < \tfrac{d-\mu}{\sigma})$$



http://www.geogebra.org/classic#probability

別説明

$X \sim \mathrm{N}(\mu, \sigma^2)$ のとき, 積分で $z = \frac{x-\mu}{\sigma}$, $\mathrm{d}z = \frac{1}{\sigma}\mathrm{d}x$ とすると,

$$
\begin{aligned}
P(c < X < d) &= \int_c^d \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x \\
&= \int_{\frac{c-\mu}{\sigma}}^{\frac{d-\mu}{\sigma}} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{z^2}{2}}\,\sigma\mathrm{d}z \\
&= P\left(\frac{c-\mu}{\sigma} < Z < \frac{d-\mu}{\sigma}\right)
\end{aligned}
$$

L08-Q4

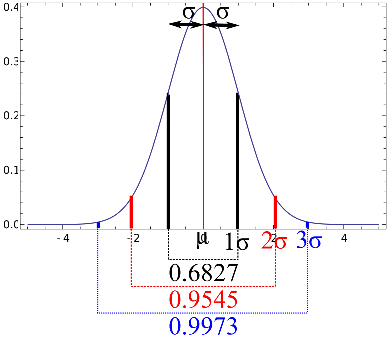
Quiz(正規分布の確率)

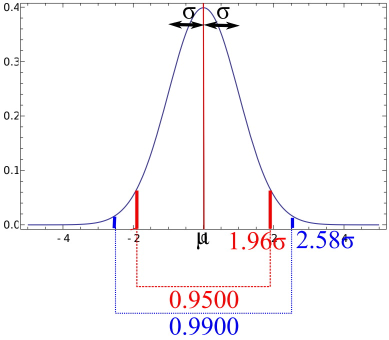
確率変数 $X \sim \mathrm{N}(3, 2^2)$ とする. 次

1. 母期待値 $\mathrm{E}[X^2]$ を求めよう.
2. $P(X \geq 5)$ を, $Q(z)$ または $I(z)$ (ただし $0 < z < +\infty$) で書こう. さらに, 表を使って小数で書こう.
3. $P(+1 \leq X \leq 7)$ を, $Q(z)$ または $I(z)$ (ただし $0 < z < +\infty$) で書こう. さらに, 表を使って小数で書こう.

岩薩林 確率・統計 §4.5 例題 4.10,4.11, 問題 10, 第 4 章練習問題 5

正規分布のだいたいの形と面積





ここまで来たよ

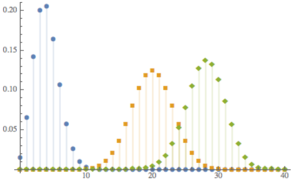
二項分布 高校 数学 B 岩薩林 確率・統計 §3.4

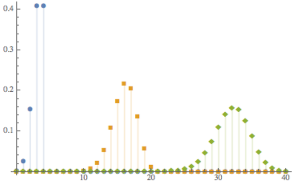
二項分布 岩薩林 確率・統計 (3.24)p.66

離散型確率変数 X が次の確率分布を持つとき, X はパラメタ n, p の二項分布 $\mathrm{B}(n,p)$ にしたがうという.

$$p(x) = \begin{cases} {}_n\mathrm{C}_x\, p^x(1-p)^{n-x} & (x = 0, 1, 2, 3, \ldots, n) \\ 0 & (\text{他}) \end{cases}$$

意味: 確率 p で表の出るコインを n 回投げたとき, x 回表が出る確率.





$${}_n\mathrm{C}_x = \binom{n}{x} = \frac{n!}{x!(n-x)!}.$$

$\mathrm{B}(40, 0.1), \mathrm{B}(40, 0.5), \mathrm{B}(40, 0.7)$, $\mathrm{B}(4, 0.8), \mathrm{B}(20, 0.8), \mathrm{B}(40, 0.8)$

二項分布の母平均値と母分散 (証明延期) 岩薩林 確率・統計定理 3.1(p.70)

$$\mathrm{E}[X] = \boxed{}, \mathrm{V}[X] = \boxed{}$$

$\mathrm{E}[1] =$

$$\boxed{}$$

二項定理 高校 数学 A 岩薩林 確率・統計 (3.25)p.68

$$(a+b)^n = \sum_{x=0}^{n} {}_n\mathrm{C}_x a^x b^{n-x}$$

L09-Q1

Quiz(二項分布)

確率 $p = \frac{2}{3}$ で表のでるコインを 100 回投げる．

1. 表が 40 回でる確率を求めよう．階乗 $n!$ とべき乗 a^b と分数 $\frac{a}{b}$ は簡単化・約分しなくてよい．P,C などの記号は使わないで答えること．
2. 表がでる回数の母平均値，母分散を求めよう．

ベルヌーイ分布

ベルヌーイ分布 岩薩林 確率・統計名前は出てこない

$n = 1$ の二項分布 $\mathrm{B}(1, p)$ のこと

$$P(X = x) = f(x) = \begin{cases} 1 - p & (x = 0) \\ p & (x = 1) \\ 0 & (\text{他}) \end{cases}$$

意味:ベルヌーイ試行=(不公平な) コイン投げ．表がでる確率 p．表 $x = 1$．

ベルヌーイ分布の母平均値と母分散

$$\mathrm{E}[X] = \boxed{}, \quad \mathrm{V}[X] = \boxed{}.$$ (上の f から直接計算すると)

ここまで来たよ

8 連続型確率変数の例:正規分布

9 二項分布, 独立同分布の和
- 二項分布
- 独立同分布
- チェビシェフの不等式と大数の法則

ベルヌーイ分布と二項分布のもうひとつの関係

岩薩林 確率・統計例題 4.6 の最初の 1 文

$X_1, X_2, \ldots, X_n$ が独立で $X_i \sim \mathrm{B}(1, p)$ のとき (独立同分布),
$T_n = X_1 + \cdots + X_n$ は $T_n \sim \mathrm{B}(n, p)$.

なぜなら

注: $T_2 = X_1 + X_2$ と, $T = 2X_1$ は異なる.

独立同分布の性質 岩薩林 確率・統計例題 4.6

独立同分布 (i.i.d.) 岩薩林 確率・統計定理 4.2(p.87) の仮定,p.113

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 (同じ確率密度関数 $f(x)$) 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.

X_1=青いサイコロ, X_2=緑のサイコロ,

(復習) 確率変数の和の母平均値と母分散

確率変数 $X_1, X_2, Y = X_1 + X_2$ を考える.
いつでも $\mathrm{E}[Y] = \mathrm{E}[aX_1 + bX_2] = a\mathrm{E}[X_1] + b\mathrm{E}[X_2]$.
X_1, X_2 が独立のとき $\mathrm{V}[Y] = \mathrm{V}[aX_1 + bX_2] = a^2\mathrm{V}[X_1] + b^2\mathrm{V}[X_2]$.

L09-Q2

Quiz(独立同分布の母期待値母分散)

確率変数 X_1, X_2, X_3 は独立同分布に従い, $\mathrm{E}[X_i] = \mu, \mathrm{V}[X_i] = \sigma^2$ とする.
次の母期待値, 母分散を μ と σ^2 で表そう.

1. $\mathrm{E}[X_1 + X_1 + X_1]$
2. $\mathrm{E}[X_1 + X_2 + X_3]$
3. $\mathrm{E}[X_1 + X_1 + X_2]$
4. $\mathrm{V}[X_1 + X_1 + X_1]$
5. $\mathrm{V}[X_1 + X_2 + X_3]$
6. $\mathrm{V}[X_1 + X_1 + X_2]$
7. $\mathrm{E}[(X_1 + X_2 + X_3)^2]$

二項分布の母平均値，母分散の式の証明 岩薩林 確率・統計例題 4.6

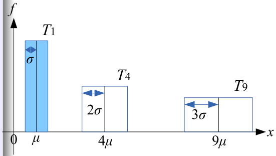
i.i.d にしたがう確率変数の和

母平均値 $\mathrm{E}[X_i] = \mu$, 母分散 $\mathrm{V}[X_i] = \sigma^2$.
和の確率変数 $T_n = X_1 + \cdots + X_n$.

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

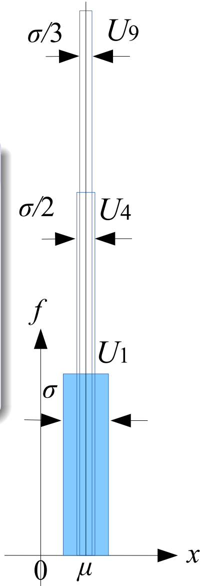
$$\mathrm{V}[T_n] = \sum_{i=1}^{n} \mathrm{V}[X_i] = n \times \sigma^2$$

T_n の確率密度関数はこんな感じ?



よって，$X \sim \mathrm{B}(n,p)$ に対して，$\mathrm{E}[X] = n \times p, \mathrm{V}[X] = n \times p(1-p)$.

U_n の確率関数は
こんな感じ？



i.i.d にしたがう確率変数の和の $1/n$

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

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

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

ここまで来たよ

チェビシェフの不等式

チェビシェフの不等式 Chebyshev's inequality 岩薩林 確率・統計 §4.3(4.15)

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

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

証明 岩薩林 確率・統計
どんな X にも使えて便利な不等式. 意味は…

L09-Q3

Quiz(独立同分布にしたがう変数の和)

確率変数 $X_1, \ldots, X_n$ は $\mathrm{E}[X_i] = \mu, \mathrm{V}[X_i] = \sigma^2$ 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．

1. 確率変数 $A = \frac{1}{n}(X_1 + X_2 + X_3 + \cdots + X_n)$
2. 確率変数 $B = \frac{1}{\sqrt{n}}(X_1 + X_2 + X_3 + \cdots + X_n - n\mu)$
3. 確率変数 $C = \frac{1}{\sqrt{n}\sigma}(X_1 + X_2 + X_3 + \cdots + X_n - n\mu)$

大数の (弱) 法則アバウト版 岩薩林 確率・統計 §4.4

$X_1, \ldots, X_n$ が独立同分布にしたがい, $\mathrm{E}[X_i] = \mu$, $\mathrm{V}[X_i] = \sigma^2$,
$U_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ のとき, n が十分大きいとき U_n は 'ほぼ μ に等しい'
(U_n が μ から外れる確率はゼロに近づく)

大数の (弱) 法則 岩薩林 確率・統計 §4.4

$X_1, \ldots, X_n$ が独立同分布にしたがい, $\mathrm{E}[X_i] = \mu$, $\mathrm{V}[X_i] = \sigma^2$,$U_n = \frac{1}{n}\sum_{i=1}^n X_i$ のとき,

$$\forall \epsilon > 0 \lim_{n \to +\infty} P(|U_n - \mu| \geq \epsilon) = 0.$$

つまり n 大で U_n は $\mathrm{E}[U_n] = \mu$ に「必ず近い」(確率収束).

証明　$\mathrm{E}[U_n] = \mu, \mathrm{V}[U_n] = \sigma^2/n$ に注意して, U_n に対するチェビシェフの不等式を書くと,

$$P(|U_n - \mu| \geq a \times \frac{\sigma}{\sqrt{n}}) \leq \frac{1}{a^2}$$

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

$$P(|U_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2 n} \to 0$$

これが母平均値・母期待値の直観的意味. 要するに,

連絡

- 2020-01-20 月 1 テスト B1 実施. 45 分, PC 受験の予定. 教科書の p. xii–xv のコピーを配布. 持込不可, 外部記憶ペーパーなし.
- 2020-01-27 月 1(定期試験期間) テスト A2,B2 実施. A1,B1 と同等の評価内容.
- 成績の 70 ピーナッツ分$=\min(\max(A1, A2), \max(B1, B2))$.
- 予習復習問題を, 期限後も (再/初) 受験できるようにしてます. ピーナッツにはカウントしないけど, テスト準備に活用してね.
- 2019-12-09 の Trial 予告
- 2019-12-09 までの (短い作文の) レポート課題 `https://manaba.ryukoku.ac.jp`
- オフィスアワー木 6(1-539) 金昼 (1-542), Math ラウンジ (1-536/538)

学籍番号での LINE 公式アカウント登録

Wolfram Alpha

龍大 manaba course

Wolfram Alpha は, スマホで定積分やグラフ描画のできる Mathematica の親戚
Google Play, AppStore に安価な有料版アプリもあります.