正規分布と中心極限定理

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 L08(2014-11-28 Fri)

今日の目標

- 正規分布の確率を求められる
- 中心極限定理の主張を説明できる

http://hig3.net

L07-S1
1,2 のみ計算できればよい.
Quiz 解答:2 項分布

1. $\sum_{k=0}^{n} {}_n\mathrm{C}_k p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$
2. $P(X < 2) = \sum_{k=0}^{\infty} p(k) \mathbf{1}_{[X<2]}(k) = p(0) + p(1) = (1-p)^n + np(1-p)^{n-1}.$
3. 実は $\mathrm{E}[X] = np.$
4. 実は $\mathrm{V}[X] = np(1-p).$
5. $\sqrt{\mathrm{V}[X]} = \sqrt{np(1-p)}.$

L07-S2
1 のみ計算できればよい.
Quiz 解答:幾何分布

1. $P(X \leq n) = \sum_{k=1}^{n} p(1-p)^{k-1} = p \cdot \frac{1-(1-p)^n}{1-(1-p)} = 1 - (1-p)^n.$ この意味は…

2. 実は $1/p$.
3. 実は $(1-p)/p^2$.

L07-S3

Quiz 解答:連続的な値をとる確率変数

1. $$\int_{-\infty}^{+\infty} f(x)\mathbf{1}_{[X \geq \frac{1}{4}]}(x)\ \mathrm{d}x = \int_{1/4}^{1/2} 8x\ \mathrm{d}x = \frac{3}{4}.$$

2. $$\mathrm{E}[X] = \int_{0}^{1/2} f(x) \cdot x\ \mathrm{d}x = 1/3.$$

3. $$\mathrm{V}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 = 1/16 - (1/6)^2 = 5/36.$$

4. $$\mathrm{E}[\frac{1}{\sqrt{X}}] = 2^{5/2}/3.$$

L07-S4
Quiz 解答:連続的な値をとる確率変数

1. $$\int_{-\infty}^{+\infty} f(x)\mathbf{1}_{[0\le X<2]}(x)\,\mathrm{d}x = \int_1^2 \frac{1}{x}\,\mathrm{d}x = \log 2.$$

2. $$\mathrm{E}[X] = \int_{-\infty}^{+\infty} f(x)\cdot x\,\mathrm{d}x = \int_1^{\mathrm{e}} \frac{x}{x}\,\mathrm{d}x = \mathrm{e}-1.$$

3. $$\mathrm{V}[X] = \mathrm{E}[X^2]-(\mathrm{E}[X])^2 = \frac{1}{2}(\mathrm{e}^2-1)-(\mathrm{e}-1)^2 = 2\mathrm{e}.$$

4. $$\mathrm{E}[\tfrac{1}{X}] = \int_{-\infty}^{+\infty} \frac{1}{x}\cdot\frac{1}{x}\,\mathrm{d}x = 1-\mathrm{e}^{-1}$$

ここまで来たよ

連続型確率変数の復習

確率密度関数 $f(x)$

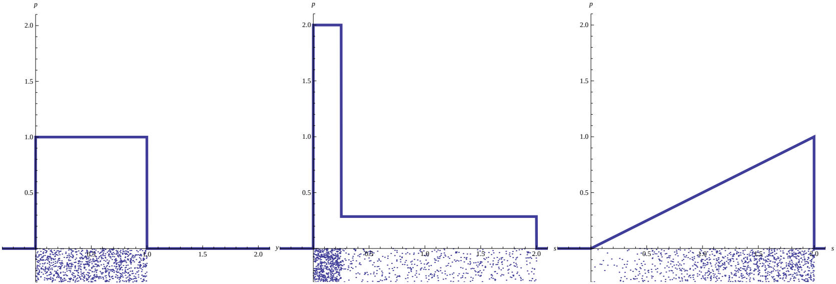
$$\varphi(x) \text{ の期待値}\quad \mathrm{E}[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x) f(x)\,\mathrm{d}x.$$

$$\text{確率 } P(a \leq X < b) = \int_a^b \mathbf{1}_{[a \leq X < b]}(x) f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x$$

累積分布関数

$$F(a) = \int_{-\infty}^{a} f(x)\mathrm{d}x = P(\boxed{\phantom{a \leq X < b}}).$$

確率密度関数の例



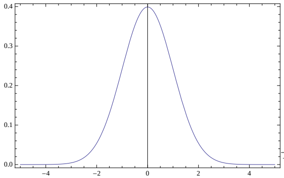
累積分布関数

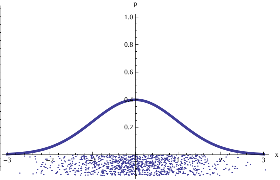
標準正規分布 (ガウス分布)

標準正規分布 $\mathrm{N}(0,1)$

$$\text{確率密度関数}\ \phi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$$

$$\text{累積分布関数}\ \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x'^2}{2}}\ \mathrm{d}x'$$



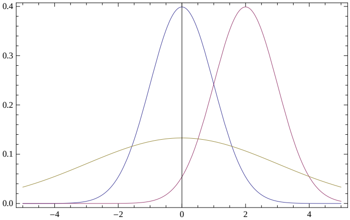


N(0, 1) の性質

$X \sim \mathrm{N}(0,1)$ のとき,

- $\mathrm{E}[1] = 1$　確率だからもちろんそうなんだけど, 確かめる計算はたいへん　微積分・演習 II
- 母平均値 $\mathrm{E}[X] = 0$
- 母分散 $\mathrm{V}[X] = 1.$

一般の正規分布 (ガウス分布)



確率密度関数

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

を考えよう

- 母平均値 $\mu = \mathrm{E}[Y] = \cdots = b$.
- 母分散 $\sigma^2 = \mathrm{V}[Y] = \cdots = a^2$.

$f(y; b, a^2)$ のグラフは, $\phi(x)$ のグラフを,

- 横に a 倍, 縦に $1/a$ 倍拡大し
- 横に b だけ平行移動した

もの ($y = ax + b$).

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

確率密度関数

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

母平均値 μ, 母分散 σ^2.

L08-Q1

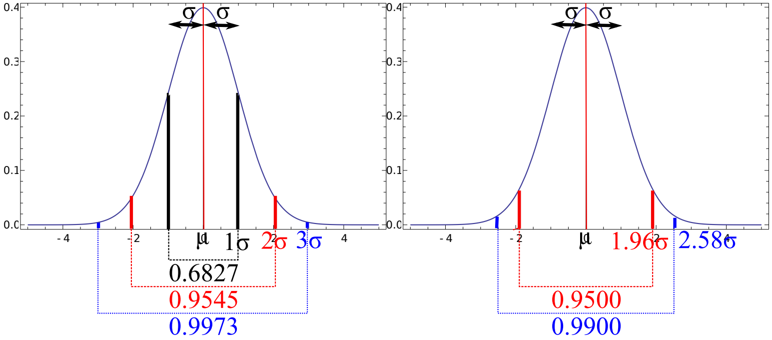
Quiz(正規分布の確率密度関数の拡大縮小平行移動)

母平均値 3, 母分散 4 の正規分布と, 標準正規分布の, 確率密度関数を重ねて描こう.

ここまで来たよ

1. 略解:連続型確率変数
2. 正規分布と中心極限定理
   - 正規分布
   - 正規分布の確率
   - 独立同分布と中心極限定理

正規分布 (ガウス分布) のグラフに関係した面積



標準正規確率表 (上側確率 $Q(x)$)

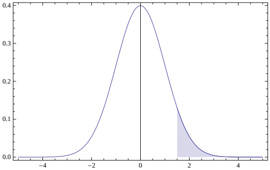
$\mathrm{N}(0,1)$ で, $X \geq x$ となる確率 $= Q(x) = 1 - \Phi(x) = \frac{1}{2}\mathrm{erfc}(x/\sqrt{2})$. Φ:累積分布関数.

紙と鉛筆では計算できない. 表またはソフトウェアに頼る.

$2(1-\Phi(1)) = \boxed{}$, $2(1-\Phi(2)) = \boxed{}$

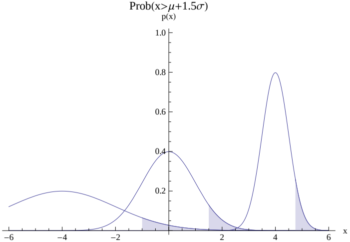
標準正規確率表 (上側確率=$Q(x) = 1 - \Phi(x)$)

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010



$\mathrm{N}(\mu, \sigma^2)$ の確率の求め方

$\mathrm{N}(0,1)$ とほとんど同じ



‘対応する’ 部分の面積は同じ

L08-Q2

Quiz(正規分布の確率)

母平均値 3, 母分散 4 の正規分布で,

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.

L08-Q3

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, $0.5 \le X \le 0.7$ 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, $0.5 \le X \le 0.7$ 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, $4.0 \le X \le 4.4$ となる確率を求めよう.

L08-Q4

Quiz(正規分布の確率)

母平均値 3, 母分散 4 の正規分布で,

1. $3 - a \leq X < 3 + a$ となる確率が 0.95 となるような a を求めよう.
2. $3 - a \leq X < 3 + a$ となる確率が 0.99 となるような a を求めよう.

L08-Q5

Quiz(正規分布の確率)

あるお店で, 琵琶湖特産瀬田シジミ 500g パックは店主の気まぐれの '時価' で販売される. 長年の調査から, その価格は, 母平均値 2000 円, 母分散 40000 円 ^{2} の正規分布に従うことがわかっている.
500g 買うためには, お金をいくら財布に入れておけば十分か答えよう. ただし, 絶対に買えるように, というと何億円あっても足りないので, 40 回に 1 回は足りなくて買えなくてもかまわないとする.

ここまで来たよ

1. 略解:連続型確率変数

2. 正規分布と中心極限定理
   - 正規分布
   - 正規分布の確率
   - 独立同分布と中心極限定理

独立同分布と中心極限定理 (定義・過程省略部分あり)

確率変数 $X_1, X_2, \ldots, X_n$ が, 独立で, すべて同じ確率分布に従う (同じ確率密度関数 $f(x)$ を持つ) とする. 正規分布でなくてよい.　独立事象 → 高校数学

これを $X_1, \ldots, X_n$ は $\boxed{}$, 独立同分布に従う, という

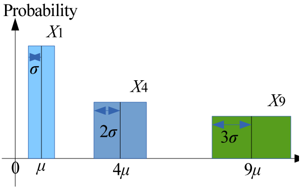
新しい確率変数を定義: $Y_n = X_1 + \cdots + X_n$.

母平均値 $\mathrm{E}[X_i] = \mu$, 母分散 $\mathrm{V}[X_i] = \sigma^2$ としたとき, 実は　過程略

Y_n の確率密度関数はこんな感じ?

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

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



中心極限定理 (いいかげんバージョン)

$X_1, \ldots, X_n$ が, 同じ確率分布に従い, 独立であるとする. 母平均値 μ, 母分散 σ^2 の独立同分布に従うとする.
このとき, $Y_n = X_1 + \cdots + X_n$ の確率分布 (累積分布関数) は, $n \to +\infty$ で

- 母平均値 $\boxed{}$
- 母分散 $\boxed{}$

の $\boxed{}$ のそれに近づいていく.

$n \to +\infty$ では分布の個性がなくなる!

連絡

- 2014-11-17 から チューターは月火水木昼 (1-614).
- 2014-12-03 水 4 数理情報学科特別講義
- $\boxed{\text{2014-12-05 金 2 教室変更するかも}}$ 7-002 でも 1-542 でもない 3 号館あたりに．ポータル経由で通知します．要注意．
- $\boxed{\text{2014-12-12 金 2 休講}}$ 　しか〜し，来年度の 3 年次必修科目 学外実習・総合演習 履修説明会．2 年生は全員出席必須．1-542.
- $\boxed{\text{いつか補講 } \times 2}$