

正規分布

樋口さぶろお <https://hig3.net>

龍谷大学 先端理工学部 数理・情報科学課程

確率統計 I L05(2024-05-20 Mon)

最終更新: Time-stamp: "2024-05-20 Mon 11:21 JST hig"

今日の目標

- 岩薩林 確率・統計 §4.5 正規分布の確率密度関数, グラフを書ける. 母期待値, 確率, 分位数を求められる.



L04-Q1

Quiz 解答: 連続型一様分布

- ① $E[X^k] = \frac{1}{d-c} \int_c^d x^k dx = \frac{1}{k+1} \frac{d^{k+1} - c^{k+1}}{d-c} = \frac{1}{k+1} (d^k + d^{k-1}c + \dots + c^k).$
- ② $E[X^1] = \frac{c+d}{2}.$
- ③ $V[X] = E[X^2] - E[X^1]^2 = \frac{(d-c)^2}{12}. \quad \sqrt{V[X]} = \frac{d-c}{\sqrt{12}} \simeq \frac{d-c}{3.5}.$

L04-Q2

Quiz 解答: 連続型一様分布の応用

- ① $X \sim U(18, 20).$
- ② $E[X^1] = 19 \text{ [mm]}.$
- ③ $P(X \geq 15) = \int_{19.5}^{\infty} f(x) dx = \int_{19.5}^{20} \frac{1}{20-18} dx = \frac{1}{4}.$
- ④ $E[6X^2] = 2168 \neq 6 \cdot 19^2 = 2166 \text{ [mm}^2\text{]}.$
- ⑤ $P(6X^2 \leq 2000) = P(X \leq (\frac{2000}{6})^{1/2}) = \int_{19}^{(2000/6)^{1/2}} \frac{1}{20-18} dx = 5\sqrt{10/3} - 9.$

$$\textcircled{6} \quad E[X^3] = 6878 \neq 19^3 = 6859 \text{ [mm}^3\text{]}.$$

L04-Q3

Quiz 解答: 連続型一様分布

- $\textcircled{1} \quad X \sim U(\sqrt{3}(-\sqrt{3}) + 5, \sqrt{3}(+\sqrt{3}) + 5)$, すなわち, $X \sim U(2, 8)$. 関数とグラフ略.
- $\textcircled{2} \quad E[X] = \frac{2+8}{2}$, または, $E[X] = \sqrt{3}E[Z] + 5 = \sqrt{3} \cdot 0 + 5$.
- $\textcircled{3} \quad V[X] = \frac{(8-2)^2}{12} = 3$, または, $V[X] = (\sqrt{3})^2 V[Z] = 3 \cdot 1$.

連続型一様分布の確率の計算方法

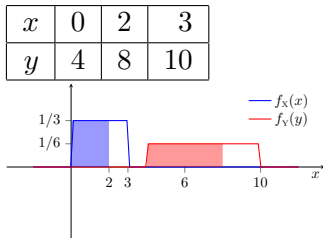
変換前 $X \sim U(0, 3)$

$P(X \leq 2) = ?$

計算方法 1 一様分布だから…

$$P(X \leq 2) = \int_{-\infty}^2 f_X(x) dx = \int_0^2 \frac{1}{3} dx = \frac{2}{3}.$$

$$f_X(x) = \begin{cases} \frac{1}{3} & (0 \leq x \leq 3) \\ 0 & (\text{他}) \end{cases}$$



変換後 $Y = 2X + 4$

$P(Y \leq 8) = ?$

計算方法 1 変換後も一様分布になるから…

$Y \sim U(4, 10)$

$$P(Y \leq 8) = \int_{-\infty}^8 f_Y(y) dy = \int_4^8 \frac{1}{6} dy = \frac{2}{3}.$$

計算方法 2 X の確率で書き直す

$$P(Y \leq 8) = P(X \leq 2) = \text{左} = \frac{2}{3}.$$

$$f_Y(y) = \begin{cases} \frac{1}{2} \cdot \frac{1}{3} & (0 \leq \frac{y-4}{2} \leq 3) \\ 0 & (\text{他}) \end{cases}$$

ここまで来たよ

4 連続型一様分布・確率変数の変数変換・標準化

5 正規分布

- 標準正規分布
- 一般の正規分布

標準正規分布 standard normal distribution $N(0, 1^2)$ の性質

定義 (標準正規分布 $N(0, 1^2)$) 岩薩林 確率・統計 (4.17)

次の確率密度関数を持つ確率変数 Z を，標準正規分布 $N(0, 1^2)$ にしたかうという。

$$f_Z(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

累積分布関数

$$F_Z(z) = \Phi(z) = \int_{-\infty}^z \phi(z') dz'.$$

ϕ, Φ は f, F のギリシャ文字．あまりに有名分布なので専用文字を使う． $\Phi(z)$ の積分は具体的に書けない．

Python で `scipy.stats.norm(loc=0, scale=1)`,

Excel で `=norm.s.dist(z, FALSE)`

標準正規分布 $Z \sim N(0, 1^2)$ の母期待値

k 次のモーメント (k : 自然数)

$$E[Z^{2k-1}] = 0, \quad \text{奇関数}$$

$$E[Z^{2k}] = (2k - 1)!!$$

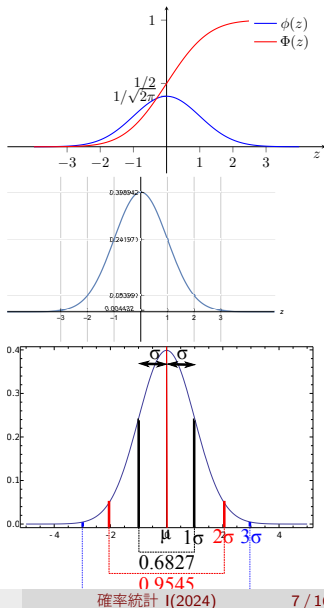
$$= 1 \cdot 3 \cdot 5 \cdots (2k - 1) \quad \text{部分積分}$$

$$\text{全確率 } E[Z^0] = 1, \quad \text{岩薩林 確率・統計 (4.19) 微積分 II}$$

$$\text{母平均値 } E[Z] = 0, \quad \text{岩薩林 確率・統計 (4.20)}$$

$$\text{母分散 } V[Z] = 1 \quad \text{岩薩林 確率・統計 (4.21)}$$

たしかに, Z は標準化された確率変数. Google Colab Python による正規分布 LearnMoodle



標準正規分布の確率と $\Phi(z)$ の数値

$Z \sim N(0, 1^2)$ のとき, 岩薩林 確率・統計 (4.22)

$$P(c < Z \leq d) = \int_c^d \phi(z') dz' = \int_c^d \frac{1}{\sqrt{2\pi}} e^{-\frac{z'^2}{2}} dz' = \Phi(d) - \Phi(c).$$

$\Phi(z)$ を `scipy.stats.norm().cdf(z)` や数表 岩薩林 確率・統計 付表 1(p.227) から求める. 分位数関数 $\Phi^{-1}(q)$ も Python や数表 岩薩林 確率・統計 付表 1(p.227) に頼る.

```

1 rvz=stats.norm(loc=0,scale=1) # Z ~ N(0, 1^2)
2 rvz.cdf(d) # Φ(d)
3 rvz.ppf(q) # Φ-1(q)

```

注: つねに $\Phi(-\infty) = 0, \Phi(+\infty) = 1$. 標準正規分布は確率密度関数が偶関数なので $\Phi(z) + \Phi(-z) = 1$. 特に $\Phi(0) = 1/2$.

岩薩林 確率・統計 付表 1 には $I(z) = \Phi(z) - \frac{1}{2}, z > 0$ の表が載っている. 高校 数学 B $z < 0$ については, $I(-z) = -I(z)$ を経由して求める.

上側確率 $Q(z) = P(Z > z) = 1 - \Phi(z)$ の表を載せてる教科書も多い.

L05-Q1

Quiz(標準正規分布の確率)

Z は標準正規分布 $N(0, 1^2)$ に従う連続型確率変数である.

- ① 確率 $P(Z \leq 1.23)$ を累積分布関数 $\Phi(z)$ で表そう. Python または表で小数として求めよう.
- ② 確率 $P(-0.56 < Z \leq +1.23)$ を累積分布関数 $\Phi(z)$ で表そう. Python または表で求めよう.
- ③ 確率 $P(Z > d) = 0.025$ となる d を累積分布関数 $\Phi(z)$ の逆関数で表そう. Python または表で求めよう.

岩薩林 確率・統計 例題 4.9(p.91), 問題 8(p.92), 問題 10(p.96)

ここまで来たよ

4 連続型一様分布・確率変数の変数変換・標準化

5 正規分布

- 標準正規分布
- 一般の正規分布

一般の正規分布 $N(b, a^2)$

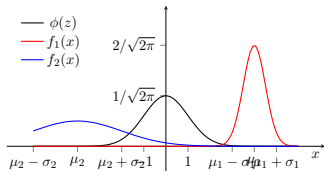
$Z \sim N(0, 1^2)$ に対して, $X = aZ + b$ を考える.

$$E[X^0] = 1,$$

$$\mu = E[X] = E[aZ + b] = b,$$

$$\sigma^2 = V[X] = V[aZ + b] = a^2,$$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \rightsquigarrow f_X(x) = \frac{1}{\sqrt{2\pi a^2}} e^{-\frac{(x-b)^2}{2a^2}}$$



定義 (一般の正規分布 $N(\mu, \sigma^2)$)

次の確率密度関数を持つ確率変数 X を, 母平均値 μ , 母分散 σ^2 の正規分布 $N(\mu, \sigma^2)$ (normal distribution) にしたがるという.

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

L05-Q2

Quiz(正規分布の確率)

連続型確率変数 X は、確率密度関数

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

にしたがう。

- 1 $E[X]$ を求めよう。
- 2 $V[X]$ を求めよう。
- 3 $f(x)$ のグラフを、標準正規分布の確率密度関数と重ねて描こう。

Python でも描いてみよう。

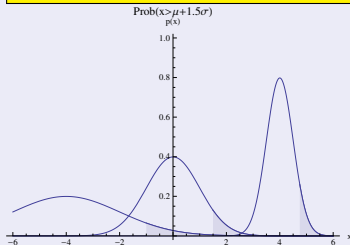
一般の正規分布 $N(\mu, \sigma^2)$ の確率

命題 (変数変換前後の確率は同じ)

確率は変数変換, 特に標準化 $Z = \frac{X-\mu}{\sigma}$ の前後で変わらない.

斜線部の面積はどれも同じ

$$\begin{aligned} P(c < X \leq d) &= P\left(\frac{c-\mu}{\sigma} < \frac{X-\mu}{\sigma} \leq \frac{d-\mu}{\sigma}\right) \\ &= P\left(\frac{c-\mu}{\sigma} < Z \leq \frac{d-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{d-\mu}{\sigma}\right) - \Phi\left(\frac{c-\mu}{\sigma}\right) \end{aligned}$$



変数				下限	上限	母平均値	母分散
z	-1	0	1	$\frac{c-\mu}{\sigma}$	$\frac{d-\mu}{\sigma}$	0	1 ²
x	$\mu - \sigma$	μ	$\mu + \sigma$	c	d	μ	σ^2

変数変換しても確率が同じことの別説明 (置換積分)

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

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

L05-Q3

Quiz(正規分布の確率)

確率変数 $X \sim N(3, 2^2)$, $Z \sim N(0, 1^2)$ とする.

- ① X, Z の確率密度関数の式を書こう. グラフを重ねて描こう.
- ② 母期待値 $E[X^2]$ を求めよう.
- ③ 確率 $P(X > 5) = P(c < Z \leq d)$ となるように c, d を定めよう. また, 確率を Python や表を使って小数で求めよう.
- ④ 確率 $P(+1 < X \leq 7) = P(c \leq Z \leq d)$ となるように c, d を定めよう. また, 確率を Python や表を使って小数で求めよう.
- ⑤ 確率 $P(+3 < X \leq 9) = P(c \leq Z \leq d)$ となるように c, d を定めよう. また, 確率を Python や表を使って小数で求めよう.
- ⑥ 確率 $P(X \leq d) = 3/4$ となる d を, 確率 $P(Z \leq d') = 3/4$ となるような d' で表そう. また, Python や表を使って小数で求めよう.

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