連続型確率変数

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　L07(2015-11-06 Fri)

最終更新: Time-stamp: "2015-11-11 Wed 07:42 JST hig"

今日の目標

- 連続型確率分布が与えられたときに, 母期待値・母平均値・母分散・母標準偏差が計算できる
- 連続型確率分布が与えられたときに, 事象の確率が計算できる

http://hig3.net

L06-Q1

Quiz 解答:共分散と相関係数

共分散 $C_{xy} = 10$

相関係数 $r = \frac{10}{\sqrt{6}\sqrt{18}} = \frac{5}{3\sqrt{3}}$.

回帰係数 $a = \frac{5}{3\sqrt{3}} \times \frac{\sqrt{18}}{\sqrt{6}} = \frac{5}{3}$.

よって回帰直線は, $y = \frac{5}{3}(x-4)+9$.

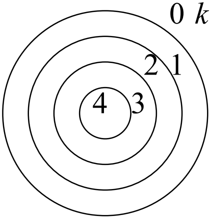
ここまで来たよ

3 回帰分析

4 連続型確率変数
- 連続型確率変数

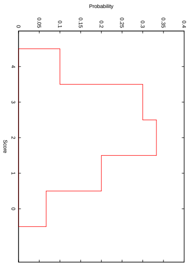
あるプレイヤーのダーツの得点確率

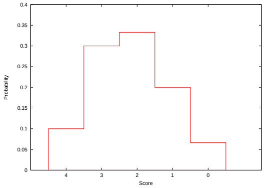
得点: 的の真ん中から順に $4, 3, 2, 1, 0$ 点



離散型確率分布

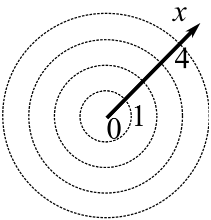
得点 x	確率 f_x
0	0.1
1	0.3
2	0.3333
3	0.2
4	0.0667

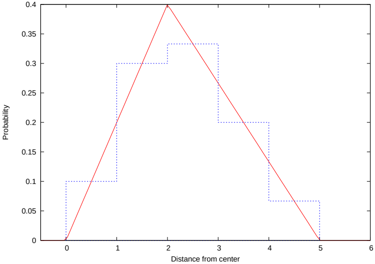




中心から x cm にあてる確率

的の真ん中からの距離 r cm, 得点 $x = 4 - r$ 点 (実数).





$r = 0.5$cm と 0.9 cm への当たりやすさは違う. $r = 1.0$cm を境に急に変わるわけじゃない. これを表現したい.

$\rightsquigarrow$ 点数の出やすさは x のある関数 $f(x)$ で表される!

連続型確率変数 連続型確率分布

確率密度関数 $f(x)$

連続型確率変数

連続型確率変数 X とは，実数値をとり，確率が確率密度関数 $f(x)$ で指定されるもの．

離散的

得点 x	確率 f_x
0	0.1
1	0.3
$\vdots$	
x	f_x
$\vdots$	

連続的

-
- $0 \leq f(x)$.
- 1 を超えることもある．

物理・工学系では $p(x)$ と書いたら確率密度関数 $f(x)$ を意味することも

連続型確率変数の母期待値

母期待値の定義

離散型確率変数 $\mathrm{E}[\phi(X)] = \sum_x f_x \cdot \phi(x)$

連続型確率変数 $\mathrm{E}[\phi(X)] = \int_{-\infty}^{+\infty} f(x) \cdot \phi(x)\, \mathrm{d}x$

$\lim_{\text{分割}\to\text{細かく}} \sum_i f(x_i)\Delta x = \int f(x)\, \mathrm{d}x$ だから自然.

- 母平均値 $\mu = \mathrm{E}[X]$, 母分散 $\mathrm{V}[X] = \mathrm{E}[(X-\mu)^2]$ など離散型と同じ定義.
- $\mathrm{E}[aX+b] = a\mathrm{E}[X]+b$, $\mathrm{V}[aX+b] = a^2\mathrm{V}[X]$, $\mathrm{V}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2$ なども成立.

母平均値, 母期待値, 母分散の直観的意味

その 1:大数の法則

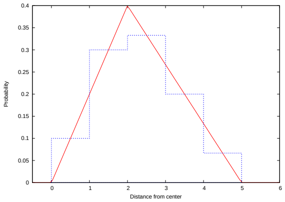
X の値 $X_1, \ldots, X_n$ を得るとき, $n \to \infty$ で,

母期待値 $\mathrm{E}[\phi(X)]$ は $T_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\phi(X_i)$ に, 「必ず近い」

正確には $\forall\epsilon > 0 \ \displaystyle\lim_{n\to+\infty} P(|T_n - \mathrm{E}[\phi(X)]| \geq \epsilon) = 0.$ (確率収束という)

要するに,

その 2:$f(x)$ のグラフ



L07-Q1

Quiz(連続的な値をとる確率変数)

次の確率密度関数を持つ確率変数 X を考える.

$$f(x) = \begin{cases} 8x & (0 \leq x < \frac{1}{2}) \\ 0 & (\text{それ以外}) \end{cases}$$

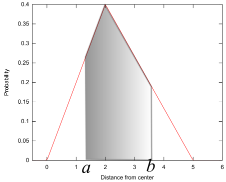
1. $X \geq +\frac{1}{4}$ となる確率を求めよう.
2. 母平均値 $\mathrm{E}[X]$ を求めよう.
3. 母分散 $\mathrm{V}[X]$ を求めよう.
4. 母期待値 $\mathrm{E}[\frac{1}{\sqrt{X}}]$ を求めよう.

確率密度関数から事象の確率を求める

$$
\begin{aligned}
P(\text{条件}) &= \mathrm{E}[\mathbf{1}_{[\text{条件}]}(X)] \\
P(a \le X < b) &= \mathrm{E}[\mathbf{1}_{[a \le X < b]}(X)] \\
&= \int_{-\infty}^{+\infty} f(x) \mathbf{1}_{[a \le X < b]}(x)\, \mathrm{d}x = \int_a^b f(x)\, \mathrm{d}x
\end{aligned}
$$

$$
\text{全事象の確率} = 1 = = \int_{-\infty}^{+\infty} f(x)\, \mathrm{d}x = \mathrm{E}[1]
$$

じゃあ, ちょうど距離 $x = a$ cm となる確率は? $\rightsquigarrow \boxed{\quad}$.



$$
\mathbf{1}_{[X \text{ の条件}]}(x) = \begin{cases} 1 & (x \text{ が条件を満たす}) \\ 0 & (\text{それ以外}) \end{cases}
$$

L07-Q2

Quiz(連続的な値をとる確率変数)

次の確率密度関数を持つ確率変数 X を考える.

$$f(x) = \begin{cases} \frac{1}{x} & (1 \leq x < \mathrm{e}) \\ 0 & (\text{それ以外}) \end{cases}$$

1. $0 \leq X < 2$ となる確率を求めよう.
2. 母平均値 $\mathrm{E}[X]$ を求めよう.
3. 母分散 $\mathrm{V}[X]$ を求めよう.
4. 母期待値 $\mathrm{E}[\frac{1}{X}]$ を求めよう.

チェビシェフの不等式

チェビシェフの不等式 Chebyshev's inequality

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.

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

どんな X にも使えて便利な不等式. 意味は…

チェビシェフの不等式の証明

$P(|X-\mu| \geq a\sigma)$ を $f(x)$ の積分で書くと…

$$
\begin{aligned}
P(|X-\mu| \geq a\sigma) &= \int_{-\infty}^{\mu-a\sigma} f(x)\,\mathrm{d}x + \int_{\mu+a\sigma}^{+\infty} \text{同じ} \\
&\leq \int_{-\infty}^{\mu-a\sigma} f(x) \times \frac{(x-\mu)^2}{(a\sigma)^2}\,\mathrm{d}x + \int_{\mu+a\sigma}^{+\infty} \text{同じ} \\
&\leq \frac{1}{(a\sigma)^2} \int_{-\infty}^{+\infty} f(x) \times (x-\mu)^2\,\mathrm{d}x \\
&= \frac{1}{(a\sigma)^2} \mathrm{V}[X] \\
&= \frac{1}{a^2}
\end{aligned}
$$

L07-Q3

Quiz(連続型確率変数の母期待値)

$\mathrm{E}[X] = 0$, $\mathrm{V}[X] = 4$ であるような連続型確率変数 X をひとつ考え, その確率密度関数 $f(x)$ を答えよう.

次回の非参照 Quiz と予習問題

- 今日の非参照 Quiz はプチテスト前に 1-503 向かいで返却．メールで連絡します．
- 来週のプチテスト，再来週 L08 には非参照 Quiz はありません
- 予習問題は 2015-11-20 金 11:05 をしめきりとして出題しますが，内容はプチテストと重なるので，プチテスト前に解いておくことをおすすめします．
- 2015-11-11 水 4,25 水 4 特別講義
- 学期半ばの授業アンケート 協力してくださる方は RaMMoodle から
- レポート課題 学期半ばのリフレクション RaMMoodle から –2015-11-13 金 23:55

連絡

- オフィスアワー月 4 木 6(1-502)

manaba 出席カード提出
`https://attend.ryukoku.ac.jp`

プチテスト計画!

- 2015-11-13 金 2, 90 分, 30 ピーナッツ, 参照相談なし．紙のテスト．
- まず授業でやらなかったページに×つけましょう．
- 過去問公開してるけどあまり参考にはならないかも．下の出題計画，非参照 Quiz, 予習問題をやり直すことをお奨めします．
- 出題計画 Excel 関係のものはありません．
  - データから平均値，分散，標準偏差を求める
  - データから (外れ値を考慮した大学レベルの) 箱ひげ図を描く
  - データから標準得点，偏差値を求める (← 注意．非参照 Quiz でカバーされてない)
  - データから共分散，相関係数を求める
  - データから回帰係数，回帰直線を求める
  - 離散型確率変数について，確率，母期待値，母平均値，母分散，母標準偏差を求める × 2 倍の量
  - 連続型確率変数について，確率，母期待値，母平均値，母分散，母標準偏差を求める × 2 倍の量 (← 注意．非参照 Quiz でまだカバーされてない)
  - 選択肢的な問