多変数の離散型確率分布

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 II L01(2016-04-14 Thu)

最終更新: Time-stamp: "2016-04-14 Thu 13:05 JST hig"

今日の目標

- 2 変数の離散型確率分布の母期待値, 周辺分布が求められる (復習) 確率統計☆演習 I(2015)L09
- 2 変数の離散型確率分布について, 同時確率, 周辺確率, 条件付き確率の間の計算ができる.

http://hig3.net

学習目標

- モーメント母関数を用いて確率分布の性質を導ける.
- 多次元分布を含む典型的な確率分布を用いて現象をモデル化し, その性質を数学的に導ける.

確率統計☆演習 II を履修してはいけない理由

次のどれも響かない人は履修しないことを奨めます.

- 中高の数学で統計はすでに強化されてる
- 教育の評価に統計は必要
- いま, 統計学が熱い! ← CPU パワー, インターネット上でのデータ集積
- いま, ビッグデータ, 人工知能 (AI), 機械学習 (machine learning) が熱い!! ← CPU パワー, インターネット上でのデータ集積
- 統計は科学技術の言葉 ⇝ 数理卒は当然期待されてる
- 確率統計を使ってる数理の教員: 松木平 (確率セルオートマトン), 馬, 佐野, 高橋 (性能評価), 飯田 (物理シミュレーション), 樋口 (確率過程, 教育評価), 他にもいるかも
- 文系でもとりあえずの技術としては統計を使える, が…
- 統計検定 2 級, 準 1 級

単位をとっているかどうかに関わらず, 確率統計☆演習 I 相当の理解があ

確率統計☆演習 II ののり

成績計算難しくないけどとにかく注文の多い科目です…
科目の成績 100 ピーナッツは

- 25 ピーナッツ:毎回授業での quiz, 授業時間外の予習復習, 授業時間内の活動など
- 30 ピーナッツ:プチテスト
- 45 ピーナッツ:ファイナルトライアル (定期試験期間)
- その他追加ピーナッツ. その時に説明.

その時点のピーナッツにかかわらず, ファイナルトライアルに参加しないと合格にはなりません. ファイナルトライアル時点で 15 ピーナッツ未満の人も, (平均点を上げるために) 参加をすすめますが, 追試験はなし.
欠席届ピーナッツ的に考慮されたい場合は, 専用用紙に事情を説明する書類を貼って, 授業前後各 5 分に提出 (事前事後とも可. ファイナルトライアルが締切). 欠席に事前連絡は原則不要. 何回欠席してもファイナルトライアル参加資格を失うことはありません.

担当者ののり

- なまえ: 樋口さぶろお hig-probstat@math.ryukoku.ac.jp
- へや: 1-502
- オフィスアワー: 月昼 (1-502), 木 4(1-502/1-539). 訪問歓迎な時間: 月木金昼 (1-502, Math ラウンジに行ってることも). お弁当持参歓迎. お湯あげます.
- Web ページ: http://hig3.net 演習の指示や, スケジュールもここから.

1 週間のタイムライン

確率統計☆演習 I(2015) とほぼ同じ (グループ座席指定以外)

1. 木 11:05 までに RaMMoodle で問題 (=非参照 Quiz 予想問題+予習問題) に解答. 次の週までは解答可能. 非参照 Quiz の 1/3 まで補充.
2. 木 2 前 先週の非参照 Quiz を返却
3. 木 2 非参照 Quiz(=テスト) 参照不可 相談不可
4. 木 2 の最後 来週の非参照 Quiz(=テスト) の予告
5. 金 09:00 予習問題公開

ここまで来たよ

1. はじめに
2. 多変数の離散型確率分布
   - 1 変数の離散型確率分布 (復習)
   - 2 変数の確率分布
   - 条件付き確率

離散的な確率変数と記号

事象の確率 P(事象). 基本事象の確率 $P(X = x) = f_X(x)$
$f(x)$:確率関数, (離散) 確率分布. 確率統計☆演習 I では f_x と書いてた.
P は事象を書く記号. $P(X^2 < 0.2)$. f はただの関数. $f_X(0.5)$.

x_i	$P(X = x_i) = f_X(x_i)$
$x_1 = 158$	$\frac{3}{6}$
$x_2 = 160$	$\frac{2}{6}$
$x_3 = 165$	$\frac{1}{6}$
合計	1

$$f_X(x) = \begin{cases} \frac{3}{6} & (x = 158) \\ \frac{2}{6} & (x = 160) \\ \frac{1}{6} & (x = 165) \\ 0 & (\text{他}) \cdots \text{省略可} \end{cases}$$

計算例 $\mathrm{E}[X] = \boxed{}$

$$= x_1 \times f_X(x_1) + x_2 \times f_X(x_2) + x_3 \times f_X(x_3)$$

$$= \sum_{i=1}^{3} x_i f(x_i) = (\text{略記}) \sum_i x_i f(x_i)$$

ここまで来たよ

1. はじめに
2. 多変数の離散型確率分布
   - 1 変数の離散型確率分布 (復習)
   - 2 変数の確率分布
   - 条件付き確率

2 次元の離散的確率変数の同時分布

同時分布

$P(X = x, Y = y) = f_{\mathrm{XY}}(x, y)$

表で書いたほうが見やすい.

$y \backslash x$	158	160	165
45	3/8	0	1/12
50	1/8	1/3	1/12

$y \backslash x$	x_1	x_2	x_3
y_1	$f_{\mathrm{XY}}(x_1, y_1)$	$f_{\mathrm{XY}}(x_2, y_1)$	$f_{\mathrm{XY}}(x_3, y_1)$
y_2	$f_{\mathrm{XY}}(x_1, y_2)$	$f_{\mathrm{XY}}(x_2, y_2)$	$f_{\mathrm{XY}}(x_3, y_2)$

計算例

$$\mathrm{E}[\tfrac{Y}{X^2}] = \boxed{}$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{2} \frac{y_j}{x_i^2} \cdot f_{\mathrm{XY}}(x_i, y_j)$$

周辺分布

周辺分布 (離散型)

$$f_X(x) = \sum_j f_{XY}(x, y_j), \quad f_Y(y) = \sum_i f_{XY}(x_i, y)$$

意味

は $\boxed{}$,

これの場合だと…

$y \backslash x$	158	160	165
45	3/8	0	1/12
50	1/8	1/3	1/12

$$f_X(x) = \begin{cases} \boxed{} \\ \boxed{} \\ \boxed{} \end{cases}, \; f_Y(y) = \begin{cases} \boxed{} \\ \boxed{} \end{cases}.$$

L01-Q1

Quiz(多次元の確率変数の期待値)

2 変数 X, Y の離散型確率分布を考える．同時分布 $f_{\mathrm{XY}}(x, y)$ が下の表で与えられる．

$y \backslash x$	1	2	3
0	0	$2/12$	$1/12$
2	$4/12$	0	$5/12$

1. 母期待値 $\mathrm{E}[X + 2Y]$ を求めよう．
2. 母期待値 $\mathrm{E}[\mathbf{1}_{[Y \geq 1]}(X, Y)]$ を求めよう．
3. 周辺分布 $f_{\mathrm{X}}(x)$, $f_{\mathrm{Y}}(y)$ を求めよう．

ここまで来たよ

1. はじめに
2. 多変数の離散型確率分布
   - 1 変数の離散型確率分布 (復習)
   - 2 変数の確率分布
   - 条件付き確率

条件付き確率

条件 B のもとでの, 事象 A の条件付き確率 $P(A|B)$
縦棒の前後は異質: P(確率を考える事象 | 条件)

意味

$y \backslash x$	158	160	165
45	3/8	0	1/12
50	1/8	1/3	1/12

条件付き確率の定義

条件 $Y = y_j$ のもとでの, $X = x_j$ の条件付き確率 $P(X = x_i|Y = y_j)$.

$$P(X = x_i|Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}, \quad f_{\mathrm{X|Y}}(x_i|y_j) = \frac{f_{\mathrm{XY}}(x_i, y_j)}{f_{\mathrm{Y}}(y_j)},$$

$$P(Y = y_j|X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)}, \quad f_{\mathrm{Y|X}}(y_j|x_i) = \frac{f_{\mathrm{XY}}(x_i, y_j)}{f_{\mathrm{X}}(x_i)}.$$

条件付き確率の性質 (1)

- 性質 0 条件付き確率

$$f_{\mathrm{X|Y}}(x|y) \neq f_{\mathrm{Y|X}}(y|x)$$

同時確率

$$f_{\mathrm{XY}}(x,y) = f_{\mathrm{YX}}(y,x)$$

- 性質 1

$$\sum_i f_{\mathrm{X|Y}}(x_i|y_j) = 1, \sum_j f_{\mathrm{Y|X}}(y_j|x_i) = 1$$

- 性質 1'

$$\sum_j f_{\mathrm{X|Y}}(x_i|y_j) \neq 1, \sum_i f_{\mathrm{Y|X}}(y_j|x_i) \neq 1$$

条件付き確率の性質 (2)

- 性質 2(同時確率との関係)

$$\begin{aligned} f_{\mathrm{XY}}(x_i, y_j) =& f_{\mathrm{X|Y}}(x_i|y_j) f_{\mathrm{Y}}(y_j) \\ =& f_{\mathrm{Y|X}}(y_j|x_i) f_{\mathrm{X}}(x_i). \end{aligned}$$

導出:

- 性質 2'(周辺確率との関係)

$$f_{\mathrm{X}}(x_i) = \sum_j f_{\mathrm{X|Y}}(x_i|y_j) f_{\mathrm{Y}}(y_j)$$

$$f_{\mathrm{Y}}(y_j) = \sum_i f_{\mathrm{Y|X}}(y_j|x_i) f_{\mathrm{X}}(x_i)$$

導出:

L01-Q2

Quiz(条件付き分布)

次の 6 枚のカードから無作為に 1 枚のカードを引く.

♡7 ♡8 ♡9 ◇8 ♠9 ♣9

$X =$ 数, $Y = 0$(赤札), 1(黒札) とすると同時分布は次のようになる.

$y \backslash x$	7	8	9	計
0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	
1	0	0	$\frac{1}{3}$	
計				

1. 9 の札が出るという条件のもとで赤札が出る, 条件付き確率を求めよう.
2. 赤札が出るという条件のもとで 9 の札が出る, 条件付き確率を求めよう.

L01-Q3

Quiz(条件付き分布)

何枚かのカードのはいった袋がある.
赤札がでる確率は $\frac{1}{3}$, 黒札がでる確率は $\frac{2}{3}$ である.
赤札がでるという条件のもとで 10 の札が出る条件付き確率は $\frac{1}{10}$,
黒札がでるという条件のもとで 10 の札が出る条件付き確率は $\frac{1}{15}$ である.

1. 赤の 10 がでる確率を求めよう.
2. 何色でもいいから 10 がでる確率を求めよう.

お知らせ

- 次回非参照 Quiz (同時確率, 周辺確率, 条件付き確率)
- 確率統計☆演習 I と同じセッティングで予習問題をやりましょう.
  http://hig3.net → RaMMoodle
  https://el.math.ryukoku.ac.jp/moodle/ → 確率統計☆演習 II(2016)

https://manaba.ryukoku.ac.jp
マイページの下の方に
manaba 出席カード提出