多次元の確率変数

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 II L01(2015-04-10 Fri)

今日の目標

- 2 次元の離散型/連続型確率変数の母期待値が計算できる.
- 2 次元の離散型/連続型確率変数の周辺分布が求められる.

http://hig3.net

ここまで来たよ

1. はじめに
   - この授業どんなのり?
2. 多次元の確率変数
   - 1 次元の確率分布
   - 2 次元の確率分布

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! $\leftarrow$ 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!! $\leftarrow$ 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 $\rightsquigarrow$ 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

統計学はこんなことに答えます

1. 高校の数学で，こういう教え方導入したら，ちょっとだけ平均点が上がった．これ効果あったって言っていいの? (Evidence-based teaching)
2. YouTube から猫の動画を見つけるアルゴリズム，こう改良して，100 個の入力画像で試したら，判定精度がちょっとあがった．これで性能あがったと結論していいの? 10000 個でやり直すべき?

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

- 30 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 (11 月)
- 40 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 20 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 6(1-539), 金 6(1-502). 訪問歓迎な時間: 火木昼. 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

模索中です.

1. 木 23:55 までに RaMMoodle で問題 (=非参照 Quiz 予想問題+予習問題) に解答
2. 金 5 非参照 Quiz(=テスト) 参照不可 相談不可
3. いつ採点返却?
4. 金 5 来週の非参照 Quiz の予告
5. 月 1 予習問題公開

ここまで来たよ

離散的な確率分布とその記号

事象の確率 $P(\text{事象})$.
基本事象の確率 $P(X=x)=f(x)$　　$f(x)$:確率関数, (離散) 確率分布.

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

$k=1,2,3=m$.

$$\rightarrow f(x)=\begin{cases}\frac{3}{6} & (x=158)\\ \frac{2}{6} & (x=160)\\ \frac{1}{6} & (x=165)\\ 0 & (\text{他})\end{cases}$$

確率統計☆演習 I(2014)L06

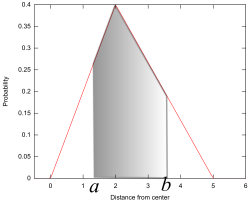
連続的な確率分布とその記号

確率密度関数の意味

$$P(a \le X < b) = \int_{-\infty}^{+\infty} f(x)\mathbf{1}_{[a \le X < b]}(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x \quad \boxed{}$$

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

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



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

確率統計☆演習 I(2014)L08

1 次元の連続的な確率変数の母期待値

期待値

離散的 $\mathrm{E}[\phi(X)] = \displaystyle\sum_{x=-\infty}^{+\infty} f(x)\cdot\phi(x),$

連続的 $\mathrm{E}[\phi(X)] = \displaystyle\int_{-\infty}^{+\infty} f(x)\cdot\phi(x)\mathrm{d}x$

ここまで来たよ

2 次元の離散的確率変数の同時分布

$P(X=x, Y=y) = f(x,y)$
表で書くと見やすい.

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

2 次元の連続的確率変数の同時分布

確率密度関数 $f(x, y)$

2 次元の確率変数の母期待値

期待値

離散的 $$\mathrm{E}[\phi(X,Y)] = \sum_{x=-\infty}^{+\infty}\sum_{y=-\infty}^{+\infty} f(x,y)\cdot\phi(x,y)$$

連続的 $$\mathrm{E}[\phi(X,Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\cdot\phi(x,y)\mathrm{d}x\,\mathrm{d}y$$

L01-Q1

Quiz(多次元の確率変数の期待値)

2 次元の連続型確率変数 (X, Y) を考える．確率密度関数 $f_{\mathrm{XY}}(x, y)$ は次で与えられる．

$$f(x,y) = \begin{cases} \frac{1}{18}xy^2 & (0 \leq x < 2, 0 \leq y < 3) \\ 0 & (\text{他}) \end{cases}$$

1. X^3Y の母期待値を求めよう．
2. $X < 1$ かつ $Y < 2$ となる確率を求めよう．

周辺分布

離散型の周辺分布

$$f_{\mathrm{X}}(x) = \sum_{y=-\infty}^{+\infty} f(x,y),$$

$$f_{\mathrm{Y}}(y) = \sum_{x=-\infty}^{+\infty} f(x,y)$$

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

連続型の周辺分布

$$
\begin{aligned}
f_{\mathrm{X}}(x) &= \int_{-\infty}^{+\infty} f(x, y)\, \mathrm{d}y \\
f_{\mathrm{Y}}(y) &= \int_{-\infty}^{+\infty} f(x, y)\, \mathrm{d}x
\end{aligned}
$$

L01-Q2

Quiz(多次元の確率変数の期待値)

2 次元の離散型確率変数 (X, Y) がある．同時確率 $P_{\mathrm{XY}}(X = x, Y = y) = f(x, y)$ が下の表で与えられる．

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

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. 周辺確率分布 $f_{\mathrm{X}}(x)$, $f_{\mathrm{Y}}(y)$ を求めよう．

L01-Q3

Quiz(多次元の確率変数の期待値)

2 次元の連続型確率変数 (X, Y) を考える．確率密度関数 $f(x, y)$ は次で与えられる．

$$f(x,y) = \begin{cases} \frac{1}{2} & (0 \leq x < 1, 0 \leq y < 1) \\ \frac{1}{4} & (2 \leq x < 3, 2 \leq y < 4) \\ 0 & (\text{他}) \end{cases}$$

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. 周辺確率密度関数 $f_{\mathrm{X}}(x)$, $f_{\mathrm{Y}}(y)$ を求めよう．

に従う