9 août 2012
PROB #16

確率論 (*Probability Theory*) 第 16 週 期末試験

w.yamamoto

はじめに：試験問題は 2 ページ以降 です。

1　諸注意とお願い

お願い：

- 解答用紙の使い方について、大問単位で、最初の１，２は同じページ、３、４はそれぞれ１ページずつ、５，６も同じページ、を使っての解答をお願いしたい。これは採点の際の都合なので、従わなくても減点はしないから、回答開始後に指示に従ってないことに気づいても、書き直さずにそのまま解答を続けて良い。

以下、どの科目でも言われる煩わしいことでしょうが、トラブル発生を避けたいので、再度、お願いしておきます。ご協力宜しくお願いします。
諸注意：

- 解答用紙は最初から 3 枚を受け取ったら、まずすべての解答用紙に名前と学籍番号を記せ。大問の解答順序は自由でよい。答案提出時には、使わなかった解答用紙も提出してよい。
- 筆記具、時計、関数電卓、衛生用品のみ、試験時間中に利用可。
- 電卓以外の機能を持つあらゆる機能、時計以外の機能を持つあらゆる機器は、持ち込みを許可しない。
- 通信手段になり得る機器、保存しておいた PDF ファイルや画像などの表示が可能な機器は、持ち込みを許可しない。(スマートフォン、タブレット、ノートパソコンなど)
- 音の出る物、振動する者は、音が出ないよう、振動もしないよう、設定を確認すること。(腕時計の時報、携帯電話の着信音とバイブレーションなど)
- 試験時間中に使用しないあらゆる物は、鞄の中にしまい、足下に置くこと。(机の中や横の座席に置いておくと、心が揺れた時や弱くなった時に、つい見たくなるかもしれないので)
- 机の中および脇の席に、何か紙が入っているだけで、ルール違反とみなして減点することがある。更にそれらに、本科目に関する記載があった場合、カンニング行為とみなす。期末試験で見つけたら、証拠物品欧州の上、その場で署名押印の書類を作成し、教務課に引き渡すことになる。
- 学生証を机の上に提示しておくこと。試験時間中に出席調査を行う。学生証もしくは仮学生証を提示できない学生については、顔写真を撮影することがある。
- ３年生は西９－１１５、２年生および４年生以上は西９－１３５で受験すること。席順は自由とする。
- 試験開始 30 分後 (03:10pm) まで遅刻者の受験を認める。退出は試験開始 30 分後 (03:10pm) から終了 5 分前 (04:05pm) まで認める。

2　期末試験問題

結果のみでなく、計算過程も読み取れるように解答すること。 数値を答える問いには、有効桁数３桁で答えよ。また必要に応じて、別紙付録の正規分布表を参照せよ。

1. 基本事項の確認。次の各項目を説明する数式、定義、もしくは表現を書け。用いる記号や変数は各自、適切に定義せよ。(各 2 点, 合計 20 点)

1-1 独立な２確率変数の分散 (の加法性): $V[X_1+X_2]=V[X_1]+V[X_2]$

1-2 独立な２確率変数の積、の期待値: $E[X_1X_2]=E[X_1]E[X_2]$

1-3 独立でない２確率変数の和、の分散: $V[X_1+X_2]=V[X_1]+V[X_2]+2Cov[X_1,X_2]$

1-4 モーメント母関数 $M_X(t)=E\left[e^{tX}\right]$

1-5 連続分布に従う確率変数、の期待値: $E[X]=\int_{-\infty}^{\infty}xf(x)\,dx$

1-6 離散分布に従う確率変数、の分散: $V[X]=E[(X-E[X])]=\sum_{k=0}^{\infty}\left(x-\sum_{j=0}^{\infty}up(u)\right)^2p(x)$

1-7 確率変数の定数倍、の分散: $V[aX]=a^2V[X]$

1-8 ２確率変数の共分散: $Cov[X_1,X_2]=E[(X_1-E[X_1])](X_2-E[X_2])$

1-9 独立とは限らない２確率変数の条件付き期待値と条件なし期待値(周辺期待値)の関係*: $E[X_1]=E\left[E_{X_1|X_2}[X_1|X_2]\right]$

1-10 独立とは限らない２確率変数の条件付き分散と条件なし分散(周辺分散とも)の関係*: $V[X_1]=E[V[X_1|X_2]]+V[E[X_1|X_2]]$

2. 不等式についての下記の問いに答えよ。ただし確率変数 X の期待値を $\mu_X=E[X]$、分散を $\sigma_X{}^2=V[X]$ と記すものとする。(各 5 点、合計 10 点)

2-1 マルコフの不等式 $Pr[X\geq a]\leq\mu_X/a$ の証明の概略を述べよ。
略解 一番短く述べるなら、「非負の確率変数の期待値 $\int_x f(x)\,dx$ が、任意の非負の定数 a について $a\int_a^{\infty}f(x)\,dx$ で下から抑えられることを用いる」だろうか。あるいは不等式

$$E[X]=\int_0^{\infty}xf(x)\,dx\geq\int_a^{\infty}xf(x)\,dx\geq a\int_a^{\infty}f(x)\,dx\geq=aPr[X\geq a] \tag{1}$$

を記しても良い。

2-2 チェビシェフの不等式 $Pr[|X-\mu_X|\geq a]\leq\sigma_X{}^2/a^2$ を用いて、$n=100$, $p=0.01$ の二項分布 $Bin(100,0.01)$ の上側 5% 点 (累積確率 95% 点) を近似的に求めよ。*
略解 二項分布について $\mu_X=np$、$\sigma_X{}^2=np(1-p)$ を既知とすれば、チェビシェフの不等式から

$$Pr[|X-np|\geq a]\leq\frac{np(1-p)}{a^2}=0.05 \tag{2}$$

という近似式を得る。期待値が 1 なため、どうせ $X<b$ の方の不等式は b が負 になるだろうから、右辺は 0.05 のまま。右辺の等式

$$a^2=np(1-p)/0.05=100\cdot0.01\cdot0.99\cdot20=19.8 \tag{3}$$

より $a=\sqrt{9.9}$ を得たら、不等式が

$$Pr\left[|X-1|\geq\sqrt{19.8}\right]\leq0.05 \tag{4}$$

となるので、

$$Pr\left[X \geq 1 + \sqrt{19.8}\right] + Pr\left[X \leq -1 - \sqrt{19.8}\right] \quad \leq \quad 0.05 \tag{5}$$

を得る。$Pr\left[X \leq -1 - \sqrt{19.8}\right]$ はもともと 0 なので、$Pr\left[X \geq 1 + \sqrt{19.8}\right] \leq 0.05$。これより、上側 5% 点を $1 + \sqrt{19.8} = 5.45$ と得るが、この点は標本空間に含まれていないので、切り上げした 6 の方が解としては望ましい。(0.10 から出発すると 4.15 あるいは 5 を得るけど、それでも正解にして。)

3. 大数の法則と中心極限定理について、下記の問いに答えよ。(各 5 点、合計 20 点)

3-1 大数の法則 $Pr\left[\left|\overline{X}_n - \mu\right| \geq \epsilon\right] \to 0 \ (n \to \infty)$ の証明の概略を述べよ。
略解 最も短い正解は「標本平均についてチェビシェフの不等式を用いる」。

3-2 大数の法則と中心極限定理の違いを述べよ。またその違いは何が主たる要因かを述べよ。
略解 最も短い解答例は「大数の法則は、標本平均が母平均に収束する、という定理であり、中心極限定理は、標本平均が母平均の回りで正規分布に従うようになる、という定理である。この違いは、前者が標本平均と母平均の距離を扱うのに対して、後者は標本平均と母平均の距離の $\sqrt{n}$ 倍を扱うところから来る。」

3-3 中心極限定理を用いて、二項分布 $Bin\,(n, p)$ の正規分布による近似を与えよ。(累積分布関数で書いても、確率分布どうしの関係で書いても、それ以外でも構わない。)
略解 あっさりと $Bin\,(n, p) \sim N\,(np, np\,(1 - p))$ でいい。

3-4 中心極限定理を用いて、$n = 100$, $p = 0.01$ の二項分布 $Bin\,(100, 0.01)$ の上側 5% 点 (累積確率 95% 点) を近似的に求めよ。
略解 2-2 と同様に。

4. 二変量正規分布についての下記の問いに答えよ。ただし $\boldsymbol{X} = (X_1, X_2)$ を二変量正規分布に従う確率ベクトルとする。また二変量正規分布の確率密度関数は、 X_1 の周辺期待値と周辺分散を μ_1 および ${\sigma_1}^2$、X_2 の周辺期待値と周辺分散を μ_2 および ${\sigma_2}^2$、X_1 と X_2 の相関係数を ρ と記すとき、

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{{\sigma_1}^2} + \frac{(x_2-\mu_2)^2}{{\sigma_2}^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right]\right\} \tag{6}$$

で与えられるものとする。(各 5 点、合計 20 点)

4-1 X_1 の周辺分布の密度関数を記せ。
略解 問題文で X_1 の周辺期待値は μ_1、周辺分散は ${\sigma_1}^2$ と教えてあり、講義中に正規分布は周辺分布も条件付き分布も正規分布と教えた上、この問題は 導け ではなく 記せ なので、これらの情報を記せば即座に

$$f_{X_1}(x_1) = \frac{1}{\sqrt{2\pi{\sigma_1}^2}} \exp\left\{-\frac{(x_1-\mu_1)^2}{2{\sigma_1}^2}\right\} \tag{7}$$

を得られるはず。

4-2 $X_1 = a$ を所与としたときの、X_2 の条件付き分布の密度関数を記せ。
略解 これは $f(a, x_2)/f(a)$ を計算すればよく、少し時間はかかるが

$$f(x_2|a) = \frac{1}{\sqrt{2\pi(1-\rho^2){\sigma_2}^2}} \exp\left\{-\frac{\left(x_2 - \left(\mu_2 + \frac{\rho\sigma_2}{\sigma_1}(a-\mu_1)\right)^2\right)}{2(1-\rho^2){\sigma_2}^2}\right\} \tag{8}$$

を得る。

4-3 相関係数 $\rho = Cov[X_1, X_2] / \sqrt{V[X_1]V[X_2]}$ が満たすべき不等式を示せ。
略解

$$-1 \leq \rho \leq 1 \tag{9}$$

4-4 4-3 で答えた不等式を証明せよ。
略解 コーシー・シュワルツの不等式を用いると

$$|Cov[X_1, X_2]| \leq \sqrt{V[X_1]V[X_2]} \tag{10}$$

が示せて、両辺を右辺で割れば

$$\rho = \frac{|Cov[X_1, X_2]|}{\sqrt{V[X_1]V[X_2]}} \leq 1 \tag{11}$$

を得る。

5. あるベルヌーイ試行 $B(p_1)$ を n 回独立に繰り返して成功回数 X_1 を観測した後で、その失敗回数分だけ別のベルヌーイ試行 $B(p_2)$ を独立に繰り返して成功回数 X_2 を観測することを考える。ここで二項分布 $Bin(n, p)$ の確率関数は

$$p(k) = \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k} \tag{12}$$

である。(各 5 点、合計 20 点)

5-1 X_1 は試行回数 n で成功確率 p_1 の二項分布 $Bin(n, p_1)$ に従う確率変数だが、X_2 は試行回数が $n - x_1$ との条件の下で成功確率 p_2 の (条件付き) 二項分布 $Bin(n - x_1, p_2)$ に従う確率変数となる。(X_1, X_2) の同時確率関数を導け。

略解 これは X_1 の周辺分布が $Bin(n, p_1)$、X_2 の条件付き分布が $Bin(n - X_1, p_2)$ と与えられているので、同時分布はこれらの積を書くだけ。ただし p_2 の意味が、X_1 で失敗したもののうちで成功する確率、という条件付き確率になっているところに注意すると

$$\begin{aligned} p(k_1, k_2) &= \binom{n}{k_1} {p_1}^{k_1} (1 - p_1)^{n-k_1} \binom{n-k_1}{k_2} p_2^{k_2} (1 - p_2)^{n-k_1-k_2} \\ &= \frac{n!}{k_1! k_2! (n - k_1 - k_2)!} p_1^{k_1} \{(1 - p_1) p_2\}^{k_2} (1 - p_1 - (1 - p_1) p_2)^{n-k_1-k_2} \end{aligned} \tag{13}$$

5-2 X_1 の周辺期待値を求めよ。

略解 周辺分布が二項分布だと書いてあるので、$E[X_1] = np_1$。

5-3 X_2 の周辺期待値を求めよ。

略解 式 (13) で、$q_1 = (1 - p_1) p_2$、$q_2 = p_1$ と置けば、X_2 の周辺分布は $Bin(n, q_1)$ となることは、5-2 より明らか。よって $E[X_2] = n(1 - p_1) p_2$。

5-4 X_1 と X_2 の共分散の符号を、理由とともに答えよ。

略解 共分散を実際に計算して $-np_1(1 - p_1) p_2$ を得た上で負と答えているか、総試行回数が n と一定で、X_1 が大きくなれば X_2 は小さくなるはずなので負と答えるか、どちらでもよい。

6. 前問にある二項分布以外で、ベルヌーイ試行から導かれる確率分布をできるだけ多く挙げ、確率分布の名前と、それぞれの確率分布とベルヌーイ試行との関係を確率分布間の関係を記せ。(最大 20 点)

解答例

1. 幾何分布：1 回成功するまでの失敗回数
2. 負の二項分布：k 回成功するまでの失敗回数
3. 超幾何分布：非復元のベルヌーイ試行の繰り返し
4. 指数分布：幾何分布から、極限分布として導かれる
5. ポアソン分布：間隔が指数分布に従う事象の一定期間の発生回数
6. 正規分布：二項分布を通じて、中心極限定理により到達

7. ギリシャ文字 (最大 20 点、別紙参照)

付録 正規分布表 (本紙は裏面が解答用紙になっており、試験後に回収 する)

(参考書より)

確率論 (J1) 2012 年度期末試験 解答用紙

学籍番号：. ______________ . 氏名：. ______________________________ .

7. 下の表を埋めなさい。(20 点、1 箇所誤るごとに 1 点減点)

7. ギリシャ文字回答欄

英語	ギリシャ大文字	ギリシャ小文字	読み (カタカナ)
a	A	α	アルファ
b	B	β	ベータ
g	Γ	γ	ガンマ
d	Δ	δ	δ
e	E	ϵ	イプシロン
z	Z	ζ	ゼータ
ê	H	η	イータ、エータ
th	Θ	θ	シータ
i	I	ι	イオタ
k	K	κ	カッパ
l	Λ	λ	ラムダ
m	M	μ	ミュー
n	N	ν	ニュー
x	Ξ	ξ	クシー、グザイ
o	O	o	オミクロン
p	Π	π	パイ
r	R	ρ	ロー
s	Σ	σ	シグマ
t	T	τ	タウ
u	Υ	υ	ウプシロン
f	Φ	ϕ	ファイ
ch	X	χ	カイ
ps	Ψ	ψ	プサイ、サイ
ô	Ω	ω	オメガ