確率統計

精選問題集



院試数学　基礎完成シリーズ

要項・類題60問

院試hub

利用上の注意

本書は院試 hub が独自に作成した基礎ノートと練習問題です。大学・研究科等の公式資料ではありません。各大学の過去問の本文・原図は収録していません。出題分野の頻度は、院試 hub が解答を公開している実施回の分類に基づきます。本書の無断転載、再配布、共有を禁じます。

目次

出題分野の頻度 5

第 1 章　確率の基礎と条件付き確率 6

確率空間と確率の基本法則 6

条件付き確率・乗法定理・ベイズの定理 6

独立性 7

問題 1　包除原理と完全順列 8

問題 2　検査の陽性的中率とベイズの定理 9

問題 3　対独立と相互独立 10

問題 4　非復元抽出と順番の対称性 11

問題 5　硬貨の種類のベイズ更新 12

問題 6　棒を折って三角形を作る確率 13

問題 7　連続する表が出るまでの回数 15

問題 8　偶数回の表と確率の漸化式 16

問題 9　ポリアの壺と交換可能性 17

問題 10　最良選択問題 18

第 2 章　確率変数・期待値・分散 20

確率変数と分布 20

期待値・分散・共分散 20

条件付き期待値 21

出題分野の頻度

院試 hub が解答を公開している実施回を分類した問数。同じ大問が複数の分野に数えられることがある。

分野	問数	主な出題校（問数）
最尤推定	29	東京大学 8、東京科学大学 7、京都大学 4、千葉大学 3、静岡大学 2
中心極限定理	16	東京大学 10、千葉大学 2、北海道大学 2、東京科学大学 1、大阪大学 1
特性関数	13	千葉大学 4、東京大学 4、東京科学大学 2、早稲田大学 1、北海道大学 1
マルコフ連鎖	7	東京大学 2、京都大学 2、東京科学大学 1、大阪大学 1、名古屋工業大学 1

のは直線的な関係の強さだけで、$E[Y \mid X]$ の形までは表さない。共分散を条件付き期待値から計算し直すと

$$\mathrm{Cov}(X,Y) = E\big[X\,E[Y \mid X]\big] - E[X]E[Y] = \frac{8\cdot\frac{1}{2} + 12\cdot\frac{5}{6}}{20} - \frac{13}{20} = \frac{14}{20} - \frac{13}{20} = \frac{1}{20}$$

となり一致する。

$V[X+Y]$ で共分散の項を落とすと $\frac{331}{400}$、2 倍し忘れると $\frac{351}{400}$ になる。分母 20 のまま計算し、最後に行和と列和の合計がそれぞれ 1 になることを確かめておくと、足し間違いも見つかる。

問題 12　三角形の台をもつ同時密度

新潟大 数理科学 2020・神戸大 経済 2026 と同じ分野

問題

(X,Y) の同時密度が

$$f(x,y) = \begin{cases} c\,x & (0 < y < x < 1) \\ 0 & (\text{その他}) \end{cases}$$

で与えられている。

(1) 定数 c を求め、X と Y の周辺密度を求めよ。

(2) $X = x$ を与えたもとでの Y の条件付き密度と $E[Y \mid X = x]$ を求めよ。

(3) $\mathrm{Cov}(X,Y)$ と $P(X+Y<1)$ を求めよ。

解答

(1)　台は $0<y<x<1$ の三角形である。x を固定すると y は $(0,x)$ を動くので

$$1 = \int_0^1\!\!\int_0^x c x\,dy\,dx = c\int_0^1 x^2 dx = \frac{c}{3} \quad\Rightarrow\quad c = 3.$$

周辺密度は

$$f_X(x) = \int_0^x 3x\,dy = 3x^2\ (0<x<1), \qquad f_Y(y) = \int_y^1 3x\,dx = \frac{3}{2}(1-y^2)\ (0<y<1).$$

(2)　$0<x<1$ について

$$f_{Y|X}(y \mid x) = \frac{f(x,y)}{f_X(x)} = \frac{3x}{3x^2} = \frac{1}{x} \quad (0<y<x)$$

なので、$X=x$ のもとで Y は $(0,x)$ 上の一様分布に従い、$E[Y \mid X=x] = \dfrac{x}{2}$ である。

(3)　$E[X] = \int_0^1 3x^3 dx = \frac{3}{4}$。全期待値の公式より $E[Y] = E[X/2] = \frac{3}{8}$。また

$$E[XY] = E\big[X\,E[Y \mid X]\big] = E\Big[\frac{X^2}{2}\Big] = \frac{1}{2}\int_0^1 3x^4 dx = \frac{3}{10}$$

だから $\mathrm{Cov}(X,Y) = \frac{3}{10} - \frac{3}{4}\cdot\frac{3}{8} = \frac{48-45}{160} = \frac{3}{160}$ である。

$P(X+Y<1)$ は台と $x+y<1$ の共通部分で積分する。$y < \min(x, 1-x)$ なので $x = \frac{1}{2}$ で分けて

$$\begin{aligned} P(X+Y<1) &= \int_0^{1/2}\!\!\int_0^x 3x\,dy\,dx + \int_{1/2}^1\!\!\int_0^{1-x} 3x\,dy\,dx \\ &= \int_0^{1/2} 3x^2 dx + \int_{1/2}^1 3x(1-x)dx = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}. \end{aligned}$$

答　(1) $c=3$、$f_X(x)=3x^2$、$f_Y(y)=\frac{3}{2}(1-y^2)$　(2) $(0,x)$ 上の一様分布、$E[Y\mid X=x]=\frac{x}{2}$　(3) $\mathrm{Cov}=\frac{3}{160}$、$P(X+Y<1)=\frac{3}{8}$

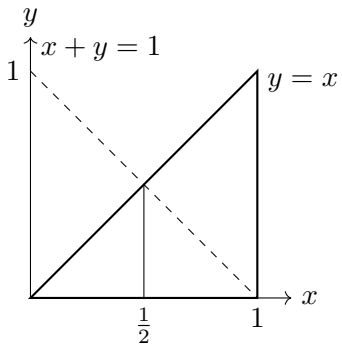

図　台（太線の三角形）と $x+y<1$ の境界（破線）。$x=\frac{1}{2}$ で上限が入れ替わる

解説

$f(x,y)=3x$ を $3x\cdot 1$ と分けて「独立」と答えるのが、この型で最も多い誤りである。台が三角形なので、例えば $x=0.1$、$y=0.5$ では $f(x,y)=0$ なのに $f_X(0.1)f_Y(0.5)>0$ となり、積に分解できない。密度の式だけでなく、台の形まで含めて判定する。

積分範囲は台の図から決める。X の周辺密度では y を 0 から x まで、Y の周辺密度では x を y から 1 まで積分する。周辺密度がそれぞれ 1 に積分されること（$\int_0^1 3x^2dx=1$、$\int_0^1 \frac{3}{2}(1-y^2)dy=1$）を確かめておく。(2) で条件付き密度が一様になるのは $f(x,y)$ が y に依存しないからで、この形が見えると $E[Y\mid X]$ は区間の中点として積分なしに書ける。周辺密度から直接 $E[Y]=\frac{3}{2}\int_0^1(y-y^3)dy=\frac{3}{8}$ と計算しても、全期待値の公式の値と一致する。

$P(X+Y<1)$ では y の上限が x と $1-x$ の小さい方になり、$x=\frac{1}{2}$ で入れ替わる。分け目を見落として $\int_0^1\int_0^{1-x}3x\,dy\,dx=\frac{1}{2}$ とすると、台の外（$y>x$）まで積分してしまう。図で破線より下にある台の部分は三角形の半分より小さく、$\frac{3}{8}<\frac{1}{2}$ と合っている。

問題 13　指示変数で数える空箱の個数

大阪公立大 数学 2025 と同じ分野

問題

n 個の玉を n 個の箱に、各玉が独立にどの箱にも等確率で入るように入れる。空の箱の個数を N とする。

(1) $E[N]$ を求めよ。

(2) $V[N]$ を求めよ。

(3) $n=4$ のときの値を求めよ。また $n\to\infty$ のとき $E[N]/n$ の極限を求めよ。

解答

箱 i が空である事象を A_i とし、指示変数 $I_i=\mathbf{1}_{A_i}$ を使うと $N=I_1+\cdots+I_n$ である。

(1)　箱 i が空であるには、n 個の玉がすべて箱 i 以外に入ればよいので $P(A_i)=\left(1-\frac{1}{n}\right)^n$。期待値の線形性より

$$E[N]=\sum_{i=1}^{n}P(A_i)=n\left(1-\frac{1}{n}\right)^n.$$

(1) の結果は、平均点で 10 点上回る Y でも、約 29% の確率で X に負けることを意味する。差の標準偏差 18 点が平均の差 10 点より大きいからである。標準化 $(D - E[D])/\sqrt{V[D]}$ の値（ここでは 0.555）を「平均の差が標準偏差の何倍か」と読むと、確率の大きさの見当がつく。

(3) は分位点の計算で、確率から点数へ逆にたどる。片側 5% の点 1.645 と両側 5% の点 1.96 を取り違えると $c = 110 + 1.96 \times 18.03 \approx 145.3$ となり、上位 2.5% の点を答えてしまう。

問題 25　一様乱数の和と積の分布

東大 数理科学 2015 と同じ分野

問題

U, V は独立にそれぞれ $(0, 1)$ 上の一様分布に従う。

(1) $S = U + V$ の密度を求め、$P(S \leq 1.5)$ を求めよ。

(2) $R = UV$ の密度を求め、$P(R \leq \frac{1}{4})$ を求めよ。

解答

(1)　独立なので畳み込みの公式より

$$f_S(s) = \int_{-\infty}^{\infty} f_U(u) f_V(s-u)\, du = \int \mathbf{1}\{0 < u < 1,\ 0 < s-u < 1\}\, du.$$

被積分関数が 1 となる u の範囲は $\max(0, s-1) < u < \min(1, s)$ で、その長さは

$$f_S(s) = \begin{cases} s & (0 < s \leq 1) \\ 2 - s & (1 < s < 2) \\ 0 & (\text{その他}) \end{cases}$$

である（三角分布）。よって

$$P(S \leq 1.5) = 1 - \int_{1.5}^{2} (2-s)\, ds = 1 - \frac{(0.5)^2}{2} = 0.875.$$

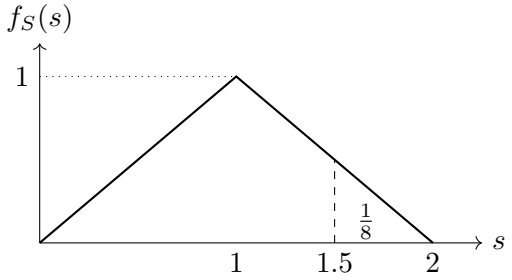

図　$U + V$ の三角形の密度。破線の右側の面積 $\frac{1}{8}$ が $P(S > 1.5)$

(2)　$0 < r < 1$ について、$U = u$ を与えると $UV \leq r \iff V \leq r/u$ である。$u \leq r$ なら確率は 1、$u > r$ なら r/u なので

$$P(R \leq r) = \int_0^r 1\, du + \int_r^1 \frac{r}{u}\, du = r - r\log r.$$

微分して $f_R(r) = 1 - \log r - 1 = -\log r$ $(0 < r < 1)$。$r = \frac{1}{4}$ を代入して

$$P\Big(R \leq \frac{1}{4}\Big) = \frac{1}{4} + \frac{1}{4}\log 4 = \frac{1}{4} + \frac{1}{2}\log 2 \approx 0.597.$$

答　(1) $f_S(s)=s$ $(0<s\le 1)$、$2-s$ $(1<s<2)$、$P(S\le 1.5)=0.875$　(2) $f_R(r)=-\log r$ $(0<r<1)$、$P(R\le \frac14)=\frac14+\frac12\log 2\approx 0.597$

解説

和の分布は、指示関数の積の積分という形にすると機械的に求まる。$f_V(s-u)=\mathbf{1}\{0<s-u<1\}$ を u の条件 $s-1<u<s$ に書き直し、f_U の条件 $0<u<1$ と重ねる。重なりの長さが s によって増えてから減るので、三角形の密度になる。図形で見るなら、正方形 $[0,1]^2$ を直線 $u+v=s$ で切った線分の長さ（を $\sqrt{2}$ で割ったもの）である。

積の分布は、和より気づきにくい形をしている。$-\log r$ は $r\to 0$ で発散し、小さい値が出やすい。二つの一様乱数を掛けると、片方が小さければ積も小さくなるからである。対数をとると $-\log R=(-\log U)+(-\log V)$ で、問題 23(1) より各項は $\mathrm{Exp}(1)$ に従うので、$-\log R\sim \mathrm{Ga}(2,1)$ となる。これを変数変換で戻すと、$P(R\le r)=P(-\log R\ge -\log r)=e^{\log r}(1-\log r)=r-r\log r$ となり (2) と一致する。和の分布を使って積の分布を求めるこの方法は、n 個の一様乱数の積にも同じ形で使え、$-\log(U_1\cdots U_n)\sim \mathrm{Ga}(n,1)$ となる。

確かめとして、$\int_0^1(-\log r)\,dr=1$、$E[R]=\int_0^1 r(-\log r)dr=\frac14=E[U]E[V]$ が成り立つ。

(1) で s の範囲を分けずに $f_S(s)=s$ を $(0,2)$ 全体に使うと、積分が 2 になる。(2) で $u\le r$ の部分（V がどんな値でも $UV\le r$）を落とすと $P(R\le r)=-r\log r$ となり、$r\to 1$ で 0 に近づくので分布関数にならない。どちらも端点での値（$F(0)=0$、$F(\text{上端})=1$）を確かめれば見つかる。

問題 26　極座標変換と 2 次元正規分布

科学大 情報工学 2025 と同じ分野

問題

X,Y は独立にそれぞれ標準正規分布 $N(0,1)$ に従う。$X=R\cos\Theta$、$Y=R\sin\Theta$（$R>0$、$0\le\Theta<2\pi$）と極座標で表す。

(1) (R,Θ) の同時密度を求め、R と Θ が独立であることを示せ。R^2 の分布も求めよ。

(2) (1) を用いて $\displaystyle\int_{-\infty}^{\infty}e^{-x^2/2}dx=\sqrt{2\pi}$ を示せ。

(3) $P(X^2+Y^2\le 4)$ と $E[R]$ を求めよ。

(4) U_1,U_2 が独立に $U(0,1)$ に従うとき、$\sqrt{-2\log U_1}\cos(2\pi U_2)$ が $N(0,1)$ に従う理由を述べよ。

解答

(1)　逆変換 $x=r\cos\theta$、$y=r\sin\theta$ のヤコビアンは

$$J=\det\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}=r$$

である。$f_{X,Y}(x,y)=\frac{1}{2\pi}e^{-(x^2+y^2)/2}$ なので、$r>0$、$0\le\theta<2\pi$ において

$$f_{R,\Theta}(r,\theta)=\frac{1}{2\pi}e^{-r^2/2}\cdot r=\left(re^{-r^2/2}\right)\cdot\frac{1}{2\pi}.$$

台が長方形 $(0,\infty)\times[0,2\pi)$ で密度が r の関数と θ の関数の積なので、R と Θ は独立で、$\Theta\sim U(0,2\pi)$、$f_R(r)=re^{-r^2/2}$ である。さらに $P(R^2>t)=P(R>\sqrt{t})=\int_{\sqrt{t}}^{\infty}re^{-r^2/2}dr=e^{-t/2}$ だから、$R^2\sim\mathrm{Exp}(\frac12)$（平均 2）である。