

「確率とその応用」ノート、その参

逆瀬川浩孝

3 条件付き確率

3.1 条件付きということ

例題 3.1 コンビニのお弁当の売り上げと天気の間係を調べるためにデータを取ってまとめたのが次の表です。このような状況が続くと考えて、以下の問に答えなさい。(1) 雨が降るという条件の下で弁当が 100 個以下しか売れない、という条件付き確率をどのように見積もればよいですか。(2) 弁当が 101 個以上売れたとして、そのとき雨だった可能性はどれくらいと見積もればよいですか（出張先で報告を受けた店長のつもりで考えて下さい）。

	100 個以下	101 個以上
雨	15 日	5 日
雨ではない	25 日	55 日

こたえ (1) 雨が降るという条件を付けた場合、過去 20 日分のデータだけしか参考ができない。そのうち 15 日は 100 個以下しか売れなかったのだから、次の雨降りの日も 100 個以下の可能性は $3/4$ くらいと考えるべきだろう。

(2) 101 個以上売れたのが 60 日あったので、その経験を元に計算することができる。結果は $5/60$ 。雨降りでも 101 個以上売れないということはないが、その可能性は非常に少ない、と思うべきである。

例題 3.2 当たりくじが 1 本しかないくじを n 人で引くとき、2 番目に引く人の当たる確率を計算しなさい。

こたえ 2 番目に引く人が当たるという事象を A 、1 番目に引く人が当たるという事象を B とすると、全確率の公式から

$$P(A) = P(A \cap B) + P(A \cap B^c) = P(A | B)P(B) + P(A | B^c)P(B^c)$$

1 番目に引く人は当たりくじ 1 本を含む n 本の中から選ぶので $P(B) = \frac{1}{n}$ 、1 番目の人が外れたとき、2 番目の人は当たりくじ 1 本を含む $n - 1$ 本の中から選ぶので $P(A | B^c) = \frac{1}{n-1}$ 、1 番目の人が当たればもちろん $P(A | B) = 0$ 、したがって、

$$P(A) = 0 \times \frac{1}{n} + \frac{1}{n-1} \times \left(1 - \frac{1}{n}\right) = \frac{1}{n}$$

3.2 ベイズの定理

例題 3.3 あるタイプのエイズ検査では、感染者は 100% 陽性と判定され、非感染者のうち 1% は陽性と判定される。エイズの感染率 0.1% としたとき、陽性判定の人が本当に感染している可能性はどれくらいあるか計算しなさい。

こたえ 全体の人口を 1 億人とする、感染している人は 10 万人、その人たちの検査結果は陽性。一方、残りの 9990 万人のうち、1% の 99.9 万人が陽性ということですので、それを次のよ

うな表を作って整理しておくことにします。

原因	その人数 $N \times P(B_i)$	因果律 $P(A B_i)$	陽性者の数
感染 B_1	10(万人)	1.0	10 (万人)
非感染 B_2	9990	0.01	99.9

全部で 109.9 万人が陽性判定されますが、そのうち感染している人は 10 万人です。したがって、陽性反応結果を受けて実際に感染している可能性は $10/109.9 = 0.09(9\%)$ ということになります。

例題 3.4 オリンピックのドーピング検査で筋力増強剤を使っていると判定されるとメダルが剥奪されます。検査では筋力増強剤を使っている人は 90% 見つかりますが、使っていない人でも 5% は陽性判定を受けるといわれています。全体の 10% 位の人は実際に使っているのではないかと噂されています。ある選手が陽性判定を受けたとして、その選手が無実でメダルを剥奪される可能性はどれくらいありますか。

こたえ 1000 人の選手集団を考えましょう。増強剤を使っている 100 人のうち 9 割の 90 人が陽性判定を受けます。使っていない 900 人の検査結果からも 45 人の陽性判定が出ます。

原因	その人数 $N \times P(B_i)$	因果律 $P(A B_i)$	陽性者の数
筋肉増強剤使用 B_1	100	0.9	90
不使用 B_2	900	0.05	45

したがって、三人に一人は身に覚えのない罪でメダルを剥奪されることになります。

例題 3.5 三つの工場から出荷されている製品に不良品が見つかった、三つの工場の製品不良率はそれぞれ 1%, 2%, 4% と言われ、また、出荷割合は 5:4:1 です。さてその製品はどの工場から出荷されたものでしょうか。

こたえ 原因は三つの工場から出荷したということ、それを B_1, B_2, B_3 とすると、出荷品総数 10000 個の場合の因果律による分類の表は、次のように構成されます。

原因	出荷個数 $N \times P(B_i)$	因果律 $P(A B_i)$	不良個数
工場 B_1	5000	0.01	50
B_2	4000	0.02	80
B_3	1000	0.04	40

工場 B_1 の製品不良率は 1% なので、出荷された製品 5000 個のうちの 50 個は不良品、というように読みます。したがって、不良品が出たとき、その製造元工場がたとえば B_2 である可能性は

$$P(B_2 | A) = \frac{80}{50 + 80 + 40} = \frac{8}{17}$$

と計算できます。

例題 3.6 ある 1 台のタクシーがひき逃げしたことが目撃された。目撃者は車体の色は青だったと言っている。この証言の信憑性が裁判の鍵になった。この町には青色と黒色のタクシーしかなくて、事故は夜だったこともあり、あるいは黒を青と誤認したのかもしれない。実際同じ条件で実験してみたところ、誤認識率は約 2 割あったという。また、当日同じ時間帯に走っていたタクシーの 85% は黒色のタクシーだった。ひき逃げしたのが本当に青色のタクシーだった可能性はどれくらいでしょうか。

こたえ これは、経済学者のトベルスキとカーネマンが奇妙な例として考え出したものです。この場合、今までの例のように、確率を人数、個数のような整数に変えるための「総数」が明確ではありませんが、同じような状況を $N = 1000$ 回経験したとしたらどうなるか、と考えます。850 回は黒のタクシーが走っているところに遭遇し、170 回はご認識して青いタクシーと判断します。一方、150 回は青のタクシーが走っているところに遭遇し、8 割の 120 回だけ正しく青のタクシーと判定します。これらを表にまとめると、次のようになるでしょう。

原因	その台数 $N \times P(B_i)$	因果律 $P(A B_i)$	青の台数
黒 B_1	850	0.2	170
青 B_2	150	0.8	120

したがって、証言のうち正しいのは半分以下の $120/(120+170)$ です。

例題 3.7 (1) A 氏は集団検診で胃がん末期と診断されて目の前が真っ暗。本当にがんなのかしら。以下のデータを参考にして検討しなさい。集団検診で末期がんと診断されたケースは過去の例から言って、実際はがんではないのがんと診断する誤診率は 5% だという。一方、本当はがんなのに見逃してしまう確率は 2% くらいと言われている。胃がんの罹患率は 10 万人当たり 1000 人である。(2) B 氏は日頃からヘビースモーカーかつ大酒のみで知られる。B 氏は最近からだがだるく感じられてしょうがない。C 医師の診断を仰いだところ肺がんだと宣告されてしまった。B 氏としても半分は自業自得とあきらめはするものの、半信半疑である。B 氏は本当にがんなのだろうか。以下のデータを参考にして考えなさい。医師は名医として知られており、がん患者を見逃すことは 2% くらい、健康な人をがんだと診断する見立て違いは 1% くらいだと言われている ((1) の考え方とどう違うのでしょうか)。

ヒント: (1) A 氏は国民からランダムに選ばれた一人として、一般に集団検診でがんと診断された場合その診断の信頼性はどうか、と考える。(2) B 氏は自分でも半分はがんかもしれないと思っているので診断を受ける前のがんであるという確率(思いこみ、自己診断)を 0.5 と考えてベイズの定理を適用する。あるいは、同じような境遇の人がたくさんいて、半分の人のがんとしたら、がんを宣告される人がそのうちの何人か(何割か)を計算すればよい。

こたえ (1) この問題での「誤診率」は A さんの立場に立ったもので、「がん」でないのに「がん」と判定する確率をいうものとします。「見逃す」というのは「(がんの人を) 見逃す」ということでしょうね、当然のことながら。1 万人が集団検診を受けたとしよう。そうすると

原因	人数 $N \times P(.)$	因果律 $P(\text{ガン診断} .)$	がん診断数
胃がん患者 B	100	0.98	98
B^c	9900	0.05	495

となるので、「がん」診断を受けた人の約 6 分の 5、83.5% は間違い!!

(2) 健康な人も含めての集団検診と違って、「がん宣告されて半分はしょうがないとあきらめる」ような人たちが 1000 人集まって同じ行動を取ったとします。その人達は半分くらいは肺がんにかかっているかもしれないと考えているとかもしれません。その前提で (1) と同じような計算をすると、

原因	人数 $N \times P(.)$	因果律 $P(\text{ガン診断} .)$	がん診断数
胃がん患者 B	500	0.98	495
B^c	500	0.01	5

となって、**絶望的!!** となります。もうちょっと楽観的になって 500:500 という数値を 100:900 にしたところで、状況はあまり改善されません。

(1) では胃がんの罹病率という統計数字が計算の根拠になっていますが、(2) ではヘビースモーカーの気持ちを数量化して計算の出発点とするという、かなりアヤシげな計算をしているところが違ってきます。ベイズの定理は、客観的なデータが得られないこのような場合にも適用可能、というよりも、そのような状況で計算できる方法として考案されました。アヤシげなので、最初は受け入れられませんでした。だんだんと認められるようになりました。「適当な」確率から出発して、計算し、「適当な」確率をどの程度動かしたらその結論が変わるのか、というような使い方をします。「適当な」確率は**主観確率**と呼ばれています。

練習問題 3.1 選択肢が 4 問ある問題に答えてもらう。答えを知らない人はランダムに解答するものとする。ある集団では解答を知っている人が 2 割しかいないとき、その集団の中からランダムに選んだ人が正解を答えたとき、その人が本当に正しい選択肢を知っていた可能性はどれくらいありますか。

練習問題 3.2 パソコンのバッテリーがすぐ壊れてしまった（充電ができない）、何とかならないか、という苦情が窓口に持ち込まれた。バッテリーは A 工場と B 工場で製造している。このバッテリーはどちらの工場から出荷された可能性が高いか。ベイズの公式を用いて計算しなさい。ただし、A 工場の出荷量は B 工場の出荷量の 3 倍、B 工場の不良率は A 工場の 2 倍である。

ヒント：バッテリーが壊れるという事象を F 、バッテリーが A 工場製であるという事象を A 、とすると $P(A | F)$ が問題。出荷量の仮定は $P(A) = 3P(B)$ と表し、不良率の仮定は $P(F | B) = 2P(F | A)$ と表すことが出来る。

練習問題 3.3 ケータイを無くしてしまった、バイト先か、居酒屋か、2 次会のカラオケか、三箇所の中のどこかに置き忘れた、ということは確かだ。この順に気が緩んでいったと思われるので、その 3 箇所に置き忘れた可能性は $1:2:4$ としよう。つまり、カラオケに置き忘れた確率はバイト先に置き忘れた確率の 4 倍、居酒屋に置き忘れた確率の 2 倍、と見積もることにします。まず置き忘れた可能性が最も大きいカラオケを探したのですが見つかりませんでした。さて、置き忘れた場所が居酒屋である可能性はどれくらいか、計算しなさい。ただし、カラオケにあったとしても、そこを探して見つかる可能性は 6 割だといわれています。

練習問題 3.4 三つの USB メモリー A, B, C を使い分けているため、せっかく集めたデータの Excel ファイルがどこにあるのか分からなくなってしまった。使用頻度から考えて、 A, B, C にある可能性は $4:2:1$ くらい。最初に A を検索したのだが、検索結果があまりたくさんあって、調べた範囲では発見することが出来なかった。発見できなかったからといって、 A にある可能性が消えたわけではなく、3 割くらいの可能性で見つけ損なった可能性がある。このとき、 B に目的のファイルがある可能性はどれくらいあるか計算しなさい。

練習問題 3.5 はしかが流行している。ある島には A, B 二つの行政区があるが、病院は一つしかない。 A, B の人口比は $3:2$ 、 A, B のはしか発生率はそれぞれ $1\%, 3\%$ としたとき、病院にきたはしか患者が A 地区から来た可能性はどれくらいか。

3.3 結合確率関数、周辺確率関数、独立性

例題 3.8 X, Y の結合確率関数が次のように与えられているものとします。このとき、以下の小問に答えなさい。(1) Y の周辺確率関数を計算しなさい。(2) $Y = 1$ という条件の下で X の

条件付き確率関数を計算しなさい。(3) X の周辺確率関数を計算しなさい。(4) X, Y は独立ですか。

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	0.1	0.1	0.05
$X = 1$	0.15	0.2	0.05
$X = 2$	0.05	0.1	0.2

こたえ (1) 定義通り

$$\begin{aligned} P(Y = 0) &= P(Y = 0, X = 0) + P(Y = 0, X = 1) + P(Y = 0, X = 2) \\ &= 0.1 + 0.15 + 0.05 = 0.3 \\ P(Y = 1) &= 0.4, P(Y = 2) = 0.3 \end{aligned}$$

(2) (1) で計算した周辺確率を使います。

$$\begin{aligned} P(X = 0 | Y = 1) &= \frac{P(X = 0, Y = 1)}{P(Y = 1)} = \frac{0.1}{0.4} = 0.25 \\ P(X = 1 | Y = 1) &= 0.5, P(X = 2 | Y = 1) = 0.25 \end{aligned}$$

(3) 定義通り

$$P(X = 0) = 0.25, P(X = 1) = 0.4, P(X = 2) = 0.35$$

(4) 条件を付けても付けなくても同じ、ということが「どんな条件」に対しても言えるかどうかを確かめる必要がある。 $P(X = 0 | Y = 1) = P(X = 0)$ は成り立つが、 $P(X = 1 | Y = 1) = 0.5 \neq P(X = 1) = 0.4$ なので、独立とは言えない（一つでも条件に合わないものがあつたらだめ）。

例題 3.9 二つのさいころを振って出た目の差の絶対値を X 、和を Y としたとき、 $Y = 4$ という条件の下で X の確率関数を求めなさい。

こたえ X の取り得る値は 0 か 2、周辺確率は $P(Y = 4) = 3/36$ なので、

$$P(X = 0 | Y = 4) = \frac{1}{3}, P(X = 2 | Y = 4) = \frac{2}{3}$$

例題 3.10 二つのさいころを振って出た目の差の絶対値を X 、和を Y としたとき X, Y の結合確率関数を求めなさい。

こたえ X の取り得る値は 0, 1, 2, 3, 4, 5、 Y の取り得る値は 2, 3, ..., 12、結合確率関数を $f_{X,Y}(i, j) = f(i, j)$ とする。 Y が偶数ならば $X = 0$ となる組み合わせは一通り、 $X = 2, 4$ になる組み合わせは、あれば二通り。 Y が奇数の場合は $X = 1, 3, 5$ で、あれば二通り、と考えて順番に表を埋めて行けばよい。 $a = \frac{1}{36}$ とすると

	$Y = 2$	3	4	5	6	7	8	9	10	11	12
$X = 0$	a	0	a	0	a	0	a	0	a	0	a
1	0	$2a$	0	$2a$	0	$2a$	0	$2a$	0	$2a$	0
2	0	0	$2a$	0	$2a$	0	$2a$	0	$2a$	0	0
3	0	0	0	$2a$	0	$2a$	0	$2a$	0	0	0
4	0	0	0	0	$2a$	0	$2a$	0	0	0	0
5	0	0	0	0	0	$2a$	0	0	0	0	0

例題 3.11 離散確率変数 X は 0, ± 1 , ± 2 で一様に分布している、つまり、確率関数が次のように与えられているものとします。

$$P(X = k) = 0.2, \quad k = -2, -1, 0, 1, 2$$

$Y = X^2$ としたとき、(1) Y の確率関数を計算しなさい。(2) X, Y の結合確率関数を求めなさい。(3) $Y = 1$ という条件の下で、 X の条件付き確率関数を計算しなさい。

こたえ (1) 0 でないものだけを書く

$$P(Y = 0) = P(X = 0) = 0.2$$

$$P(Y = 1) = P(X = -1 \text{ または } X = 1) = 0.4$$

$$P(Y = 4) = P(X = -2 \text{ または } X = 2) = 0.4$$

(2) 0 でないものだけを書く

$$P(X = -2, Y = 4) = P(X = -1, Y = 1) = P(X = 0, Y = 0) = 0.2$$

$$P(X = 1, Y = 1) = P(X = 2, Y = 4) = 0.2$$

(3) 0 でないものだけを書く

$$P(X = -1 | Y = 1) = P(X = 1 | Y = 1) = \frac{1}{2}$$

例題 3.12 1 番から 4 番までの通し番号が書かれた四つの箱がある。 k 番の箱には 1 から k までの番号の付いた k 個の玉が入っている。最初に一つの箱をランダムに選び、次いで選んだ箱の中からランダムに玉を一つ選ぶ。その玉の番号を X 、最初に選んだ箱の番号を Y としたとき、(1) $P(X = j)$ を計算するために Y を使った全確率の公式を書きなさい。(2) $P(X = 2)$ を計算しなさい。

こたえ (1)

$$P(X = 4) = P(X = 4, Y = 4)$$

$$P(X = 3) = P(X = 3, Y = 3) + P(X = 3, Y = 4)$$

$$P(X = 2) = P(X = 2, Y = 2) + P(X = 2, Y = 3) + P(X = 2, Y = 4)$$

$$P(X = 1) = P(X = 1, Y = 1) + \cdots + P(X = 1, Y = 4)$$

(2) Y は 1, 2, 3, 4 という値を取る確率がそれぞれ $\frac{1}{4}$ 。 $Y = k$ という条件の下で、 X は 1, 2, ..., k という値を取る確率が $\frac{1}{k}$ 、したがって、条件付き確率で書き換えると、

$$P(X = 2, Y = 2) = P(X = 2 | Y = 2)P(Y = 2) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

$$P(X = 2, Y = 3) = P(X = 2 | Y = 3)P(Y = 3) = \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{12}$$

$$P(X = 2, Y = 4) = P(X = 2 | Y = 4)P(Y = 4) = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$$

これらを合わせて、

$$P(X = 2) = \frac{1}{8} + \frac{1}{12} + \frac{1}{16} = \frac{13}{48}$$

練習問題 3.6 $P(A | B) = P(A)$ が成り立つならば、 $P(A | B^c) = P(A)$ も成り立つことを証明しなさい。

練習問題 3.7 X, Y の結合確率関数が下の表のように与えられているものとします ($P(X = 0, Y = 0) = 0.05$ というように読む)。このとき、以下の問いに答えなさい。(1) $P(Y = k)$ ($k = 0, 1, 2$) を計算しなさい。(2) $X = 0$ という条件の下で、 Y の条件付き確率関数を計算しなさい。

(3) X と Y は独立ですか。理由を付けて答えなさい。

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	0.05	0.1	0.25
$X = 1$	0.1	0.2	0.3

練習問題 3.8 表の出る確率が p のコインを 3 回投げるとき、最初の 2 回で出る表の回数を X 、最後の 2 回で出る表の回数を Y としたとき、(1) X, Y の結合確率関数を求めなさい。(2) $Y = 1$ という条件の下で X の条件付き確率関数を求めなさい。(3) X の周辺確率関数を求めなさい。

例題 3.13 コイン投げを繰り返すという実験で、 X_k は k 回目のコイン投げで表が出たら 1、裏が出たら 0 とする確率変数として、 $S_n = X_1 + X_2 + \dots + X_n$ と定義します。このとき、以下の問いに答えなさい。ただし、「 $S_3 = 1$ 」は「 S_3 が 1 に等しいという事象」、「 $S_3 = 2, X_1 = 1$ 」は「 $S_3 = 2$ かつ $X_1 = 1$ という事象」「 $X_1 = 1 \mid S_3 = 1$ 」は「 $S_3 = 1$ という条件の下での $X_1 = 1$ という事象」を表すものとします。(1) 事象 $\{X_1 = 1 \mid S_3 = 1\}$ をコイン投げの言葉を使って説明し、その確率を求めなさい。(2) 事象 $\{X_2 = 1 \mid S_3 = 2, X_1 = 1\}$ をコイン投げの言葉を使って説明し、その確率を求めなさい。(3) $P(X_2 = 1 \mid S_3 = 1)$ 、 $P(X_2 = 1 \mid S_3 = 2, X_1 = 0)$ 、 $P(X_2 = 1 \mid S_3 = 2)$ を求めなさい。

こたえ (1) 3 回投げて 1 回表が出た、ということが分かっているとき、表が最初に出たという事象、その確率は、3 回のうちいつ出てもおかしくないで $\frac{1}{3}$ 。

(2) 3 回投げて 2 回表、そのうちの 1 回は最初、ということが分かっているとき、2 回目も表だったという事象、2 回目か 3 回目どちらかが表で、その確率は等しいので $\frac{1}{2}$ 。

(3) 同様に考えて

$$\begin{aligned} P(X_2 = 1 \mid S_3 = 1) &= \frac{1}{3}, P(X_2 = 1 \mid S_3 = 2, X_1 = 0) = 1 \\ P(X_2 = 1 \mid S_3 = 2) &= P(X_1 = 0, X_2 = 1 \mid S_3 = 2) \\ &\quad + P(X_1 = X_2 = 1 \mid S_3 = 2) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3} \end{aligned}$$

最後の式は全確率の公式を使っている。

練習問題 3.9 赤いさいころを振ったときに出る目を X 、黒いさいころを振ったときに出る目の数を Y として、 $Z = \max(X, Y)$ 、 $W = \min(X, Y)$ と定義します。このとき、(1) Z の確率関数を求めなさい。(2) W の確率関数を求めなさい。(3) Z と W の結合確率関数を求めなさい。(4) Z, W は互いに独立ですか。

ヒント： X の値を使って全確率の公式を適用しなさい。あるいは（上級者のためのヒント） $(Z \leq k) \Leftrightarrow (X \leq k, Y \leq k)$ という関係を使いなさい。

練習問題 3.10 (続き) 同じ記号を使って、(1) $Z = 4$ という条件の下で X の条件付き確率関数を求めなさい。(2) X と Z は独立ですか。

練習問題 3.11 (続き) 同じ記号を使って、

$$P(X + Y = k) = \sum_j P(X + Y = k \mid Y = j)P(Y = j)$$

という関係を使い、 X, Y が互いに独立であることを利用して、 $X + Y$ の確率関数を計算しなさい。

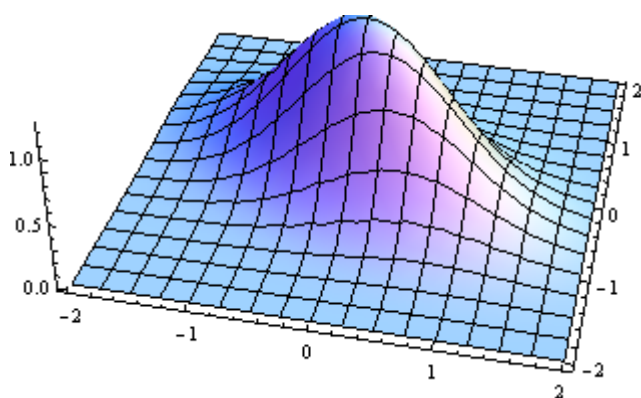
ヒント：分からないときは可能な k の値をすべてリストアップし、順番に X, Y の取りうる値の範囲を調べなさい。

例 3.1 ダーツボードでダーツの刺さった位置を直角座標系で (X, Y) と描くと、デタラメに投げた場合の X, Y の結合密度関数は、ダーツボード上で一様（一定値）になります。ダーツボードを単位正方形 $[0, 1] \times [0, 1]$ と表せば、 X, Y の結合密度関数は次のように表されます。

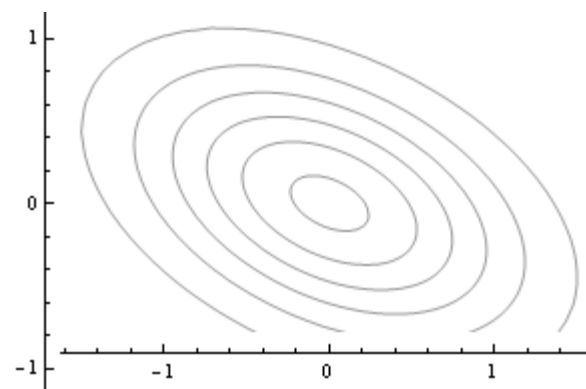
$$f_{X,Y}(x,y) = \begin{cases} 1, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{それ以外} \end{cases}$$

例 3.2 もし、ダーツボードに標的が描いてあるとすると、中心に行くほど刺さる可能性が高くなるでしょうから、その場合の結合密度関数は富士山のような形状、すなわち、**2 変量正規分布**の結合密度関数のようになるはずです。その等高線は、標的のような同心円になるでしょう。

例 3.3 X を日本人成年男子の身長、 Y を同じく体重として X, Y の確率分布を考えると、上の 2 次元ヒストグラムのように、両者の平均値を中心に富士山のような単峰の図形が得られるでしょう。測定値に丸めなければ、結合密度関数のようななめらかな曲面が予想されます。その理想的な曲面が、教科書 59 ページにある 2 変量正規分布の密度関数です。



2 変量正規分布の密度関数



2 変量正規分布の密度関数（等高線）

例題 3.14 2 変量正規分布の密度関数

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} e^{-\xi(x,y)/(2(1-\rho^2))}$$

$$\xi(x,y) = \left(\frac{x-\mu}{\sigma}\right)^2 - 2\rho\frac{x-\mu}{\sigma}\frac{y-\nu}{\tau} + \left(\frac{y-\nu}{\tau}\right)^2$$

を x について積分することにより、 Y の周辺密度関数を計算しなさい。

こたえ μ （ミュー）、 ν （ニュー）、 σ （シグマ）、 τ （タウ）。 Y の周辺密度関数は、 X, Y の結合密度関数を x について積分することによって求めることができます。

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} e^{-\xi(x,y)/2/(1-\rho^2)} dx$$

ここで、記号を整理するために、 $u = (x - \mu)/\sigma, v = (y - \nu)/\tau$ という変数を導入すると、

$\xi(x, y) = u^2 + v^2 - 2\rho uv$, $\frac{dx}{\sigma} = du$ なので、

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} \frac{1}{2\pi\tau\sqrt{1-\rho^2}} \exp\left(-\frac{u^2 + v^2 - 2\rho uv}{2(1-\rho^2)}\right) du \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi\tau\sqrt{1-\rho^2}} \exp\left(-\frac{(u-\rho v)^2}{2(1-\rho^2)} - \frac{v^2}{2}\right) du \\ &= \frac{1}{\sqrt{2\pi}\tau} e^{-v^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{(u-\rho v)^2}{2(1-\rho^2)}\right) du \end{aligned}$$

最後に得られた式の積分の被積分関数は、パラメータ $\rho v, \sqrt{1-\rho^2}$ の正規分布の密度関数であることに注意すると、密度関数の性質から全域での積分は 1 になります。残ったものに $v = (y - \nu)/\tau$ を代入すると、

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\tau} \exp\left(-\frac{(y-\nu)^2}{2\tau^2}\right)$$

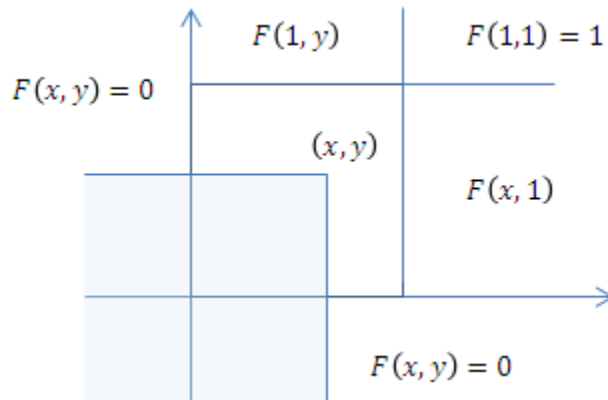
という、パラメータ ν, τ^2 の正規分布の密度関数が得られます。

練習問題 3.12 X, Y がパラメータ $(\mu, \sigma^2), (\nu, \tau^2), \rho$ の 2 変量正規分布に従うとき、 $Y = y$ という条件の下での X の条件付き密度関数が、平均 $\mu + \rho\sigma(y - \nu)/\tau$ 、分散 $\sigma^2(1 - \rho^2)$ の正規分布にしたがうことを示しなさい。

例題 3.15 X, Y の結合密度関数が次のように与えられているとき、その結合累積分布関数を計算しなさい。

$$f_{X,Y}(x, y) = \begin{cases} 1, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{それ以外} \end{cases}$$

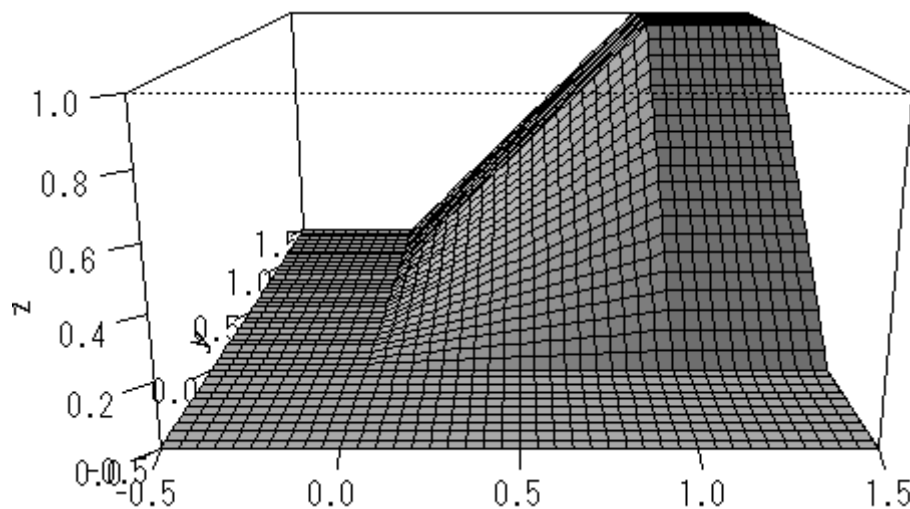
こたえ 累積分布関数 $F(x, y)$ は、密度関数の等高線図で点 (x, y) の左下の部分、の体積を計算すればよい。例えば $F(0.5, 0.5) = 0.25, F(1, 0.3) = 0.3$ のように。



この問題では $x < 0$ 、あるいは $y < 0$ ならば $F(x, y) = 0$ 。 $x < 1$ で $y > 1$ ならば $F(x, 1)$ と同じ。 $x > 1$ で $y < 1$ ならば $F(x, y) = 1$ などが分かる。したがって、計算を要するのは $0 < x, y < 1$ の場合だけ

$$F(x, y) = \int_0^x \int_0^y dx dy = xy, \quad 0 < x, y < 1$$

その 3-d グラフは次のようになる。 $y = x$ の断面で放物線になる。



練習問題 3.13 ラーメン屋のカウンターは8席、カウンターは入り口から奥へ一列になっている。開店後最初の客が一番手前の椅子を除いた7席の中からランダムに選んだ一つの椅子に座る。2番目の客は最初の客が座った席より入り口に近い席をランダムに一つ選んで座る。このとき、2人の客の間に二つ以上空き席がある確率はいくつですか。

ヒント：1番目の客が座った位置によって条件を付けて各条件ごとに考えなさい。1番目の客が2番目か3番目の椅子に座った場合は確率0です。4番目の椅子に座った場合はどうなりますか、5番目の場合は？6番目の場合は？

練習問題 3.14 (ラーメン屋の連続バージョン) 確率変数 Y は区間 $[0, 1]$ で一様分布、確率変数 X は区間 $[0, Y]$ で一様分布するものとします。このとき、**(1)** $Y = b$ の条件の下での X の条件付き累積分布関数 $P(X \leq x | Y = b)$ を計算し、 x で微分することにより、条件付き密度関数 $f_{X|Y}(x|b)$ を求めなさい。**(2)** (X, Y) の結合密度関数 $f_{X,Y}(x, y)$ を求めなさい。**(3)** これを使って、 $Y - X > \frac{1}{3}$ となる確率を計算しなさい。

ヒント：確率変数が連続の場合は、数え上げ、という考え方が出来ないなので、直感が働かないかもしれないが、考え方は離散の場合と同じ。教科書 61, 62 ページを復習してください。(1) は離散の場合と同じように考えればよい。一様分布については教科書 46 ページを復習してください。(2) は Y が一様分布に従うことから、同じく教科書 46 ページの知識を使ってください。(3) は $Y - X > \frac{1}{3}$ を満たす (X, Y) を直交座標に表してみれば積分領域が分かる。 $\int \int_{y-x > \frac{1}{3}} f(x, y) dx dy$ を計算すればよい。