数理工学モデル化実習

— 新聞売り子の抱える問題からポートフォリオまで—

担当：山本芳嗣・竹原浩太

2011 年

1　悲しいコータ君（1 時間目）

彼はコータ君という名前の新聞売り子である。彼は新聞（H 新聞）を 1 部につき 50 円で仕入れて、100 円で売り、売れ残った分は翌日 1 部 20 円で引き取ってもらっている。コータ君は売れ行きのよい日を予測することはできなかったので、ときどき損をする日もあった。以下は、この経験が彼を統計的解析に導いてくれた話である。

2　記録するコータ君

コータ君はあまり商売にならなかった日の晩よく考え、翌日から販売日誌をつけることにした。40 部以上売れた日は 1 日もなく、35 部以上売れた日はごく希にある。また 30 部以上売れた日はいくらかあり、20 部あるいはそれ以上売れた日はしばしばあった。なお売れ行きの良い日にはお客さんに新聞を売るのを断ったこともあるのを思い出した。

さらに、コータ君は 10 部単位ごとに仕入額と売上額を計算し、表 1 に可能な利潤をまとめた。仕入れ部数を示す横の行と、需要部数を示す縦の列との交差する場所に利潤がわかるようになっている。

この表を見てコータ君は途方に暮れた。

> 《もしも 50 部仕入れて 50 部売れると利潤は 2,500 円だ。だけど、もし全部売れ残ると 1,500 円も損をする。20 部仕入れて 20 部売れると利潤は 1,000 円だけど、全部売れ残ると 600 円の損。損をしないのは 1 部も仕入れないときだけ！》

Warming-up：仕入れ部数を 35 にしたとき、0, 10, 20, 30, 40, 50 の需要に対応する利潤はいくらか？

表 1: 仕入れ部数と需要部数と利潤（単位：円）の関係

		需要部数					
		0	10	20	30	40	50
仕入れ部数	0	0	0	0	0	0	0
	10	-300	500	500	500	500	500
	20	-600	200	1,000	1,000	1,000	1,000
	30	-900	-100	700	1,500	1,500	1,500
	40	-1,200	-400	400	1,200	2,000	2,000
	50	-1,500	-700	100	900	1,700	2,500

h

表 2: コータ君が観測した新聞の 1 日あたり需要分布表

需要部数	13	14	15	16	17	18	19	20	21	22	23	24	25
その需要のあった日数	1	1	1	0	5	5	8	4	2	5	7	8	4

需要部数	26	27	28	29	30	31	32	33	34	35	36	37
その需要のあった日数	7	8	10	8	5	3	2	3	1	0	1	1

表 3: コータ君が観測した 1 日あたり需要分布表（10 部毎バージョン）

需要部数の階級	0-4	5-14	15-24	25-34	35-44	45-	合計
階級値	0	10	20	30	40	-	-
その需要のあった日数	0	2	45	51	2	0	100
対応する確率	0.00	0.02	0.45	0.51	0.02	0.00	1.00

3　さらに記録するコータ君

事業の成功は、常によい観察に裏付けられている。困ったあげく、コータ君さらによく毎日の現象をつかむことにした。

> 《大事なのはその日その日の利潤ではなく、1 カ月、2 カ月あるいはもっと長期間にどれだけ利潤をあげることができるかということだ。毎日同じ部数の新聞を仕入れることにしたら、かなり長い期間にわたっての利潤を計算できるはずだ。そこで必要なのは、この仕入れ部数を決めることと、お客さんの買い方を知ることだ。よし！これからしばらくの間、僕は新聞が売り切れても売場を離れないで、まだ売れ残った新聞がある振りをして、毎晩 7 時まで立って、需要部数を記録してやろう。もし新聞が売り切れていたら、お客さんには「すみません」って謝ればいいんだから。》

こうして 100 日間の記録が始まった。それをまとめたのが表 2 である。

4　統計学者への道

彼は続けてこう言っている。

> 《もし将来が過去と同じパターンを持つと考えられるとして、僕が毎日同じ部数の新聞を仕入れると、どんな結果になるかが分かったら便利だ。この度数分布表を使えば、それぞれの仕入れ部数に応じて、100 日間の総利潤を計算できる。それを 100 で割れば、1 日あたりの平均利潤が求められる。》

ここでコータ君はこの経験的に得られた確率を用いて、仕入れ部数ごとの見込まれる平均利潤を計算した。ただし、ここでは計算の手間を少なくするため、表 2 の需要分布を 10 部単位の需要分布に丸めた表 3 を用いて、近似的な計算をしている。彼の計算の詳細と結果は、表 4 に示してある。

> 《結局、1 日の仕入れ部数を ____ 部にするといい。そのとき、平均利潤は ____ 円になる。》

と彼は結論する。

表 4: 見込まれる平均利潤

仕入れ部数	見込まれる平均利潤
0	=0
10	-300(0.00) + 500(0.02) + 500(0.45) +500(0.51) +500(0.02)+500(0.00)=500
20	-600(0.00) + 200(0.02) + 1,000(0.45) +1,000(0.51) +1,000(0.02)+1,000(0.00)=984
30	
40	
50	

$\langle\langle$ **メモ** $\rangle\rangle$ **期待値、分散、標準偏差、変動係数**　より一般的に、確率変数 D のとり得る値を $d_1, ..., d_n$, それぞれが起こる確率を $p_1, ..., p_n$ とすると、その期待値 $\mathrm{E}[D]$ は

$$\mathrm{E}[D] := \sum_{i=1}^{n} p_i \, d_i$$

と定義される。期待値 $\mathrm{E}[D]$ は D として平均的にどんな値をとるかを示していると考えられる。

例）　サイコロを 1 回振ったときの目は

$$D := \begin{cases} 1\ (= d_1) & \text{確率：}1/6\ (= p_1) \\ 2\ (= d_2) & \text{確率：}1/6\ (= p_2) \\ 3\ (= d_3) & \text{確率：}1/6\ (= p_3) \\ 4\ (= d_4) & \text{確率：}1/6\ (= p_4) \\ 5\ (= d_5) & \text{確率：}1/6\ (= p_5) \\ 6\ (= d_6) & \text{確率：}1/6\ (= p_6) \end{cases}$$

なる確率変数と考えることができ、その期待値は

$$\mathrm{E}[D] = \frac{1}{6}\cdot 1 + \frac{1}{6}\cdot 2 + \frac{1}{6}\cdot 3 + \frac{1}{6}\cdot 4 + \frac{1}{6}\cdot 5 + \frac{1}{6}\cdot 6 = 3.5$$

である。一方、確率変数 D の分散 $\mathrm{V}[D]$ は

$$\mathrm{V}[D] := \mathrm{E}\left[(D - \mathrm{E}[D])^2\right]$$

と定義される。上の流儀にならって書けば、

$$\mathrm{V}[D] = \sum_{i=1}^{n} p_i \left(d_i - \bar{d}\right)^2 = \sum_{i=1}^{n} p_i \, {d_i}^2 - \bar{d}^2$$

となる。ただし、$\bar{d} := \mathrm{E}[D]$ である。$\left(d_i - \bar{d}\right)^2$ は d_i が平均値 $\mathrm{E}[D]$ $(= \bar{d})$ から離れていればいるほど大きくなる確率変数とみなせるから、分散は確率変数が平均値 $\mathrm{E}[D]$ から平均的にどれくらいばらついているかをみる尺度と言える。特に、平方根をとって次元（単位）をもとの確率変数の次元と等しくしたものを標準偏差という：

$$（\text{標準偏差}）:= \sqrt{\mathrm{V}[D]} = \sqrt{\sum_{i=1}^{n} p_i \, (d_i - \bar{d})^2}.$$

また、次元（単位）を消したばらつきの尺度に変動係数がある：

$$（\text{変動係数}）:= \frac{（\text{標準偏差}）}{（\text{平均値}）} = \frac{\sqrt{\mathrm{V}[D]}}{\mathrm{E}[D]}.$$

5　とりあえず、やってみよう

5.1　実験 I（新聞需要のシミュレーション）

用具：サイコロ、方眼紙、紙コップ

10 個のサイコロをそれぞれ 1 回ずつ振り、それぞれのサイコロの目から 1 ずつ引いた値の合計（つまり、サイコロ目の合計値から 10 を引いたもの）を計算することを 1 回の試行とする。この試行を 40 回行い、頻度データを記録しなさい。

6　統計学者へ（2時間目）

ここから、コータ君の友人の社会工学類の学生ヨッシーとの地道な努力が始まる。ある晩 2 人は松見公園横のカフェ「パルケ」で会った。

ヨッシーは、10 部ごとのクラスにまとめて作成した表 3 ではなく、より詳しい表 2 を眺めていた。そして需要頻度を元に、その累積頻度を計算して表 5 を作った。図 1 のグラフは累積頻度の値を、需要部数の関数として描いたものである。ヨッシーは次のように言った。

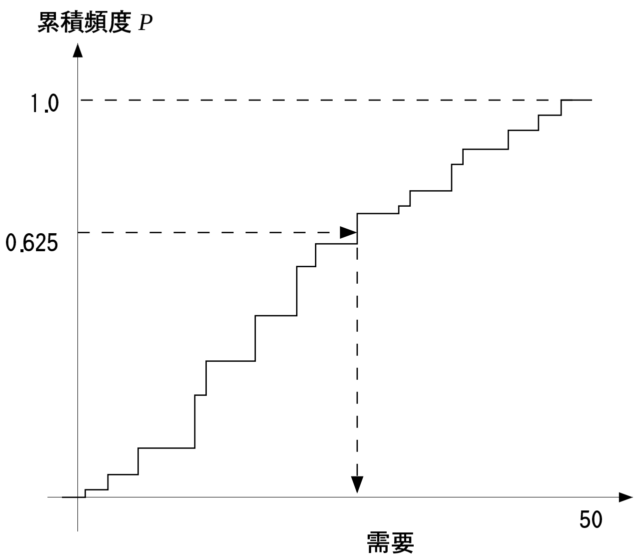

図 1: 累積頻度のグラフ

> 「もしこれからも今までと同じ形の頻度で繰り返されるんやったら、100 回の内、新聞が最大 17 部まで売れるチャンスは 8 回、18 部まで売れるチャンスは 13 回あるはずや。同んなじように、最大 20 部まで売れるチャンスは 25 回ある。これを確率と呼ぶことにしよう。いま $P(x)$ を最大 x まで売れる需要の確率とすると、もしも将来も過去と同一の度数が起きるんやったら、ものすごい簡単な論理で、最大利潤を獲得できる新聞部数を簡単に求めることができるんや。」

以下はその方法の説明である。

1 日の新聞仕入れ部数を仮に $(s-1)$ として、もう 1 部余分に仕入れていたらどうなっていたかを考えよう。もし、需要が $(s-1)$ 部より多かったら、50 円だけ利潤が多くなるはずだが、その確率は $1-P(s-1)$

表 5: 累積頻度表

需要	頻度	累積頻度
$<$10	0	0
10	0	0
11	0	0
12	0	0
13	1	1
14	1	2
15	1	3
16	0	3
17	5	8
18	5	13
19	8	21
20	4	25
21	2	
22	5	
23	7	
24	8	
25	4	
26	7	
27	8	
28	10	
29	8	
30	5	
31	3	92
32	2	
33	3	
34	1	
35	0	
36	1	
37	1	
38	0	
39	0	
40	0	
$>$40	0	

である。逆に需要が $(s-1)$ 部以下であったら 30 円の損失を受けるが、その確率は $P(s-1)$ である。したがって追加利潤は、

$$50[1-P(s-1)]-30P(s-1)$$

すなわち、

$$50-80P(s-1)$$

ということになる。だから、もう 1 部仕入れ部数を増やした方が得であるためには、

$$50-80P(s-1)>0$$

すなわち、

$$P(s-1)<\frac{50}{80}=0.625$$

でなければならない。だから s の値は

$$P(s-1)<0.625$$

を満たすようにしなければならない。また、この条件を満たしていれば 1 部の追加は利潤の増加をもたらす。つまり、最適な仕入れ部数、それを s^* であらわすと、その値は、

$$P(s^*-1)<0.625\leq P(s^*)$$

を満足しなければならない。そこで先の累積頻度の曲線（図 1）を見ると、$s=27$ とすれば、

$$P(26)=0.58,\quad P(27)=0.66$$

であることが分かる。従って、$s^*=27$ が最適仕入れ部数である。

27 部を毎日仕入れていたならば、平均利潤はいったいどのくらいになるだろうか？この計算は、考えられるすべての需要についての確率を使えばできる。表 6 はそうしたやり方で、1 日あたりの平均利潤 $v(27)$ の計算を行ったものである。

表 6: 平均利潤の計算

需要	見込み利潤
0	$0.0\times(-30\times 27+50\times 0)=0$
1	$0.0\times(-30\times 26+50\times 1)=0$
$\cdots$	$\cdots$
$\cdots$	$\cdots$
13	$0.01\times(-30\times 14+50\times 13)=2.3$
14	$0.01\times(-30\times 13+50\times 14)=3.1$
$\cdots$	$\cdots$
$\cdots$	$\cdots$
50	$0.0\times(-30\times 0+50\times 27)=0$
	合計 $v(s)=1,086$

しかし、$v(s)$ の計算は次の漸化式によって行うこともできる。

$$\begin{aligned} v(s) &= v(s-1)+50(1-P(s-1))-30P(s-1) \\ &= v(s-1)+50-80P(s-1) \end{aligned}$$

$v(0)$ は明らかにゼロである。以下、

$$
\begin{aligned}
v(1) &= 0 + 50 - (80)(0) = 50 \\
v(2) &= 50 + 50 - (80)(0) = 100 \\
v(3) &= 100 + 50 - (80)(0) = 150 \\
&\vdots \\
v(13) &= 600 + 50 - (80)(0.00) = 650.0 \\
v(14) &= 650 + 50 - (80)(0.01) = 699.2 \\
&\vdots
\end{aligned}
$$

となり、同様に計算を続けると、

$$
\begin{aligned}
v(26) &= 1082.4 \\
v(27) &= 1086.0 \\
v(28) &= 1083.2 \\
v(29) &= 1072.4
\end{aligned}
$$

となる。

ヨッシーは次のように言った。

> 「コータの最も良い計画は、27 部の新聞を仕入れることや。そうすると最大平均利潤、つまり 1 日当たり 1,086 円を得ることができる。ただし特別な事件が起こった日は別やで。まあ、そんなことはめったに起こらへんから、通常の日に仕事の基調をあわせておくのが安全やと思う。ほな、またな。」

7　賢くなったから、もういちどやってみよう

7.1　実験 II

1. 実験 I の 40 回のデータについて、累積頻度を求めなさい。
2. 最適仕入れ部数 s^* を上の手順で計算しなさい。
3. そのときの平均利潤 $v(s^*)$ を計算しなさい。

8　リスクを恐れるコータ君 ～平均値の落とし穴～（3時間目）

こうしてコータ君はその後 1 年間、毎日 27 部ずつ仕入れて店頭に立った。その結果、平均的には当初目論んでいた利潤を達成できていることがわかった。でも

《そもそも毎日必ず 27 部売れる訳ではないな。15 部しか売れないときもあるし、40 部近く需要のある日もある。もう少し安定した利潤が欲しいなあ。》

事情を聞くとヨッシーは

「なるほど、リスクを減らしたいんや。」

と言うと、コータ君が持参した新聞の需要データを自分のパソコンに入力し、たちどころに次の表 7 を作成した。この表は仕入れ部数に対応して既に計算した期待利潤と、利潤の標準偏差を示したものである。標

表 7: 仕入れ部数ごとの期待利潤と標準偏差

仕入れ部数	0	5	10	15	20	25	30	35	40	45	50
期待利潤	0.0	250.0	500.0	747.6	959.2	1073.2	1055.2	929.2	781.6	631.6	481.6
標準偏差	0.0	0.0	0.0	17.7	100.2	245.6	362.8	398.8	404.1	404.1	404.1

準偏差は利潤が平均値（期待利潤）からどれくらいばらついているかを示す指標の 1 つである。前節では期待利潤が最大になる仕入れ部数を定めた。表 7 からも 27 部に近い 25 部で期待利潤が大きくなっていることが分かる。一方、標準偏差の方は仕入れ部数を少なくすればするほど小さい値を示している。仕入れ部数が 10 部以下では標準偏差が 0 であるから、期待利潤分の利潤が確実に得られることが分かる。このように、通常、期待利潤を最大にすることと標準偏差を最小にすることは両立できるとは限らない。

「もしコータが標準偏差を小さくしたいんやったら、仕入れ部数を下げて期待利潤も下げんとあかんな。逆に、高い期待利潤を求めるならリスクを取らんと。」

9　もう 1 つの新聞も考えるコータ君 ～ポートフォリオ選択理論への道～

新聞売り子はコータ君だけではない。ギョーカイでは先輩格のウシオさんもやはりコータ君と同じ問題意識を持っていた。ウシオさんはいいと思ったことは直ちに実行する人である。実際、コータ君が需要を記録し始めたときにウシオさんも記録を始めた。日ごろから何かにつけお世話になっていたこともあり、コータ君はヨッシーに教わったことをウシオさんに教えていた。《リターンを求めるならリスクを取らなければ。》どこかで聞いた台詞を今度はコータ君が口にしていた。

ただ状況が違っていたのは、ウシオさんは D スポーツ紙（以下、D スポ）を専門に扱っていた点にある。D スポは仕入れ値も引き取り値もコータ君の扱う H 新聞と同じ価格である。

コータ君は早速習得した知識を活用してみたくて、ウシオさんがとった D スポの需要記録の分析をかって出た。表 8 はリターンとリスクを表にしたものである。

これを見てコータ君は愕然とした。

《たとえば、D スポを 25 部仕入れたとする。表から分かるのは、このとき平均利潤は 814.0 円で、H 新聞を 25 部仕入れたときの平均利潤の 1,073.2 円より小さいし、標準偏差は 453.6 円で、H 新聞の 245.6 円より大きい。つまり、D スポを扱うことは H 新聞を扱うよりもリターンとリスクのどちらの点から見ても不利に見える。D スポを取り扱うのは損なのだろうか。ウシオさんになんて言おうか。》

表 8: 「D スポーツ」紙の仕入れ部数ごとの期待利潤と標準偏差

仕入れ部数	0	5	10	15	20	25	30	35	40	45	50
期待利潤	0.0	250.0	480.0	673.2	798.4	814.0	711.2	573.2	423.2	273.2	123.2
標準偏差	0.0	0.0	77.9	193.0	328.6	453.6	507.2	528.6	528.6	528.6	528.6

困ったときにはヨッシーの出番である。ヨッシーはコータ君が作成した 2 つの表（表 7 と表 8）を眺めた後、H 新聞と D スポの需要記録を見せるようコータ君に言った。《既にリスクとリターンは計算してあるのに》と、コータ君がやや不満を感じながらノートを差し出すと、ヨッシーは先日入力した H 新聞の需要記録データの隣の列に、同一の日付の記録が同じ高さに並び合うように D スポの需要記録データを入力した。表 9 がそれである。

「負の相関が高いから、これはチャンスかも。」

そして、コータ君の方に向き直り、

「1 時間くれれば、期待利潤を大きく、かつ標準偏差を小さくする仕入れ部数を計算したげるわ。」
《リターンを取るにはリスクを取らなければならないのでは？》

？マーク点灯のコータ君を尻目にパソコンで作業を開始したヨッシーは、30 分も経たないうちに次の大きな表 10 を作成し、コータ君に説明を始めた。

「この表はやね、行は H 新聞の仕入れ部数、列は D スポーツの仕入れ部数に対応していて、表の中身はその仕入れ部数の組み合わせでの期待利潤と標準偏差を表したものなんや。」

《なるほど。でも、それって 2 つの仕入部数について足した値が出てくるだけと違うの？》

「期待利潤についてはその通りや。そやけど、標準偏差についてはそうとは限らへん。定義に立ち返って考えてみるとすぐにわかる。」

$$\begin{aligned}
\mathrm{V}[X+Y] &= \mathrm{E}[\{X+Y-(\mu_x+\mu_y)\}^2] = \mathrm{E}[\{(X-\mu_x)+(Y-\mu_y)\}^2] \\
&= \mathrm{E}[(X-\mu_x)^2 + 2(X-\mu_x)(Y-\mu_y) + (Y-\mu_y)^2] \\
&= \mathrm{V}[X] + \mathrm{V}[Y] + 2\mathrm{E}[(X-\mu_x)(Y-\mu_y)]
\end{aligned}$$

コータ君の予想は

$$\mathrm{V}[X+Y] = \mathrm{V}[X] + \mathrm{V}[Y]$$

であったことに注意していただきたい。明らかに $X+Y$ の分散と X の分散と Y の分散の和は $\mathrm{E}[(X-\mu_x)(Y-\mu_y)]$ の値（この値は共分散と呼ばれる。詳しくはメモ参照）だけ異なっている。コータ君の予想が正しいのはは $\mathrm{E}[(X-\mu_x)(Y-\mu_y)] = 0$ のときだけである。

例えば両紙併せて 25 部仕入れる場合を考えてみよう。表の太字で書かれた所を見てほしい。H 新聞のみ 25 部仕入れる場合、期待利潤は 1073.2、標準偏差は 245.6 であるのに対し、H 新聞を 15 部、D スポを 10 部仕入れた場合、期待利潤が 1227.6、標準偏差が 79.3 となり、期待利潤をより大きく、標準偏差をより小さくできることが分かる。つまり適切に組み合わせれば、単独で販売するより期待利潤（リターン）を大きく、かつ標準偏差（リスク）を小さくできる可能性が出てくる。

このことからヨッシーがアドバイスしたことは、コータ君とウシオさんはそれぞれの新聞だけを販売するのではなく、適切に 2 紙を組み合わせれば、期待利潤を変えることなく、標準偏差を下げることができるということであった。

表 9: H 新聞と D スポの需要部数記録

日付	H 新聞	D スポ	日付	H 新聞	D スポ
4 月 1 日	13	34	5 月 21 日	27	26
4 月 2 日	19	27	5 月 22 日	33	8
4 月 3 日	30	15	5 月 23 日	31	16
4 月 4 日	22	27	5 月 24 日	27	14
4 月 5 日	27	24	5 月 25 日	30	14
4 月 6 日	19	26	5 月 26 日	26	26
4 月 7 日	24	14	5 月 27 日	25	21
4 月 8 日	30	10	5 月 28 日	26	26
4 月 9 日	28	21	5 月 29 日	22	24
4 月 10 日	19	24	5 月 30 日	17	28
4 月 11 日	22	19	5 月 31 日	26	23
4 月 12 日	25	14	6 月 1 日	28	20
4 月 13 日	29	25	6 月 2 日	26	15
4 月 14 日	20	27	6 月 3 日	32	12
4 月 15 日	21	32	6 月 4 日	29	18
4 月 16 日	29	10	6 月 5 日	23	23
4 月 17 日	20	26	6 月 6 日	27	22
4 月 18 日	24	19	6 月 7 日	20	24
4 月 19 日	29	7	6 月 8 日	36	14
4 月 20 日	29	17	6 月 9 日	26	26
4 月 21 日	28	19	6 月 10 日	22	25
4 月 22 日	17	19	6 月 11 日	27	16
4 月 23 日	18	25	6 月 12 日	25	20
4 月 24 日	25	29	6 月 13 日	29	17
4 月 25 日	28	21	6 月 14 日	32	16
4 月 26 日	24	26	6 月 15 日	33	10
4 月 27 日	33	12	6 月 16 日	30	8
4 月 28 日	19	24	6 月 17 日	27	10
4 月 29 日	31	15	6 月 18 日	19	27
4 月 30 日	14	35	6 月 19 日	27	7
5 月 1 日	24	17	6 月 20 日	29	17
5 月 2 日	23	16	6 月 21 日	18	27
5 月 3 日	28	23	6 月 22 日	30	16
5 月 4 日	26	19	6 月 23 日	17	28
5 月 5 日	34	5	6 月 24 日	15	18
5 月 6 日	23	21	6 月 25 日	26	20
5 月 7 日	19	32	6 月 26 日	18	27
5 月 8 日	23	26	6 月 27 日	28	21
5 月 9 日	17	22	6 月 28 日	18	21
5 月 10 日	17	24	6 月 29 日	28	10
5 月 11 日	21	25	6 月 30 日	19	22
5 月 12 日	24	22	7 月 1 日	28	18
5 月 13 日	22	22	7 月 2 日	24	22
5 月 14 日	28	19	7 月 3 日	18	27
5 月 15 日	23	24	7 月 4 日	24	23
5 月 16 日	37	5	7 月 5 日	20	32
5 月 17 日	19	23	7 月 6 日	23	26
5 月 18 日	29	20	7 月 7 日	28	5
5 月 19 日	24	18	7 月 8 日	31	20
5 月 20 日	27	26	7 月 9 日	23	21

表 10: 仕入れ部数ごとの期待利潤と標準偏差

期待利潤		D スポの仕入れ部数										
		0	5	10	15	20	25	30	35	40	45	50
	0	0.0	250.0	480.0	673.2	798.4	**814.0**	711.2	573.2	423.2	273.2	123.2
H	5	250.0	500.0	730.0	923.2	**1048.4**	1064.0	961.2	823.2	673.2	523.2	373.2
新	10	500.0	750.0	980.0	**1,173.2**	1,298.4	1,314.0	1,211.2	1,073.2	923.2	773.2	623.2
聞	15	747.6	997.6	**1,227.6**	1,420.8	1,546.0	1,561.6	1,458.8	1,320.8	1,170.8	1,020.8	870.8
の	20	959.2	**1,209.2**	1,439.2	1,632.4	1,757.6	1,773.2	1,670.4	1,532.4	1,382.4	1,232.4	1,082.4
仕	25	**1,073.2**	1,323.2	1,553.2	1,746.4	1,871.6	1,887.2	1,784.4	1,646.4	1,496.4	1,346.4	1,196.4
入	30	1,055.2	1,305.2	1,535.2	1,728.4	1,853.6	1,869.2	1,766.4	1,628.4	1,478.4	1,328.4	1,178.4
れ	35	929.2	1,179.2	1,409.2	1,602.4	1,727.6	1,743.2	1,640.4	1,502.4	1,352.4	1,202.4	1,052.4
部	40	781.6	1,031.6	1,261.6	1,454.8	1,580.0	1,595.6	1,492.8	1,354.8	1,204.8	1,054.8	904.8
数	45	631.6	881.6	1,111.6	1,304.8	1,430.0	1,445.6	1,342.8	1,204.8	1,054.8	904.8	754.8
	50	481.6	731.6	961.6	1,154.8	1,280.0	1,295.6	1,192.8	1,054.8	904.8	754.8	604.8

標準偏差		D スポの仕入れ部数										
		0	5	10	15	20	25	30	35	40	45	50
	0	0.0	0.0	77.9	193.0	328.6	**453.6**	507.2	528.6	528.6	528.6	528.6
H	5	0.0	0.0	77.9	193.0	**328.6**	453.6	507.2	528.6	528.6	528.6	528.6
新	10	0.0	0.0	77.9	**193.0**	328.6	453.6	507.2	528.6	528.6	528.6	528.6
聞	15	17.7	17.7	**79.3**	192.8	327.6	451.6	503.8	523.8	523.8	523.8	523.8
の	20	100.2	**100.2**	120.3	202.5	321.3	436.4	482.1	497.9	497.9	497.9	497.9
仕	25	**245.6**	245.6	243.6	265.6	321.8	393.4	424.3	435.2	435.2	435.2	435.2
入	30	362.8	362.8	350.3	335.9	333.1	360.2	378.7	386.8	386.8	386.8	386.8
れ	35	398.8	398.8	380.2	353.9	333.4	349.3	365.3	372.9	372.9	372.9	372.9
部	40	404.1	404.1	384.2	356.6	334.4	349.1	364.7	372.3	372.3	372.3	372.3
数	45	404.1	404.1	384.2	356.6	334.4	349.1	364.7	372.3	372.3	372.3	372.3
	50	404.1	404.1	384.2	356.6	334.4	349.1	364.7	372.3	372.3	372.3	372.3

⟨⟨ **メモ** ⟩⟩ **共分散と相関係数**　2 つの確率変数のペア (X, Y) のとり得る値を $(x_1, y_1), ..., (x_n, y_n)$, それぞれが起こる確率を $p_1, ..., p_n$ とすると、X と Y の共分散 $\mathrm{Cov}[X, Y]$ は

$$\mathrm{Cov}[X, Y] := \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])] = \sum_{i=1}^{n} p_i (x_i - \bar{x})(y_i - \bar{y})$$

と定義される。ただし、$\bar{x} := \mathrm{E}[X],\ \bar{y} := \mathrm{E}[Y]$ とおいた。以下の変形も容易に示すことができる：

$$\mathrm{Cov}[X, Y] = \mathrm{E}[X\,Y] - \mathrm{E}[X]\,\mathrm{E}[Y] = \sum_{i=1}^{n} p_i\, x_i\, y_i - \bar{x}\,\bar{y}.$$

また、X と Y の相関係数 $\rho[X, Y]$ は以下のように定義される：

$$\rho[X, Y] := \frac{\mathrm{Cov}[X, Y]}{\sqrt{\mathrm{V}[X]}\sqrt{\mathrm{V}[Y]}} = \frac{\sum\limits_{i=1}^{n} p_i\, x_i\, y_i - \bar{x}\,\bar{y}}{\sqrt{\sum\limits_{i=1}^{n} p_i\, {x_i}^2 - \bar{x}^2}\sqrt{\sum\limits_{i=1}^{n} p_i\, {y_i}^2 - \bar{y}^2}}.$$

特に、$p_i = \frac{1}{n}\ \ (i = 1, ..., n)$ ならば、

$$\hat{X} := (x_1 - \bar{x},\, x_2 - \bar{x},\, ...,\, x_n - \bar{x}), \quad \hat{Y} := (y_1 - \bar{y},\, y_2 - \bar{y},\, ...,\, y_n - \bar{y})$$

とすれば、

$$\rho[X, Y] = \frac{\langle \hat{X} \cdot \hat{Y} \rangle}{\|\hat{X}\|\,\|\hat{Y}\|} = \frac{\sum\limits_{i=1}^{n} x_i\, y_i - n\,\bar{x}\,\bar{y}}{\sqrt{\sum\limits_{i=1}^{n} {x_i}^2 - n\,\bar{x}^2}\,\sqrt{\sum\limits_{i=1}^{n} {y_i}^2 - n\,\bar{y}^2}}.$$

ここで、$\langle \cdot, \cdot \rangle$ は内積を、$\| \cdot \|$ はユークリッドノルムを示す。$\rho[X, Y]$ は -1 と 1 の間の値である。

10　最後にもう1つやってみよう

10.1　実験 III

1. 実験 I の 40 個のデータを前半 20 個と後半 20 個で 2 等分し、それぞれを $h_i,\ d_i,\ (i=1,...,20)$ とする。表 11 を利用して、それぞれの平均、分散（または標準偏差)、および、互いの相関係数を求めなさい。

2. 上で選んだ 2 つのデータそれぞれを同じ日の H 新聞と D スポの需要と見なし、

   (a) H 新聞だけ 60 部

   (b) D スポだけ 60 部

   (c) H 新聞 30 部+D スポ 30 部

   仕入れた場合の 3 つの場合について、利潤の系列を計算しなさい（末尾の表 12 の $f(h)$ と $f(d)$ の欄を埋めなさい)

3. 上の (a),(b),(c) それぞれの場合について、期待利潤と分散 （または標準偏差）を求めなさい。

4. 以下の手続きに従い、上で選んだ $h_i,\ d_i\ (i=1,...,20)$ を変換した新しい系列 $h_i',\ d_i',\ (i=1,...,20)$ を作成し、これについて上記 1.〜3. を繰り返しなさい（末尾の表を使用すると簡単）:

   (a) H 新聞と D スポの需要 $h_i,\ d_i$（$i=1,\cdots,20$）を隣り合わせに並べる。

   (b) 20 日分のデータを（1〜5 日目、6〜10 日目、11〜15 日目、16〜20 日目という様に、）4 分割する。

   (c) それぞれの分割に対して、H 新聞のデータを昇順（小さい順）に、D スポのデータを降順（大きい順）に並べ替えたものを $h_i',\ d_i'\ (i=1,...,20)$ とする。

5. 1.〜3. と 4. の結果を比較検討しなさい。

6. ［思考実験］今度は 20 日分のデータを（1〜10 日目、11〜20 日目という様に、）2 分割した上で、それぞれの分割に対して、H 新聞のデータを昇順（小さい順）に、D スポのデータを降順（大きい順）に並べ替えたものを $h_i'',\ d_i''\ (i=1,...,20)$ とした場合、どのような結果が予想されるか？なぜそう考えるのか？

参考文献

[1] 脇本和昌『統計学　見方・考え方』 日本評論社、1984 年

[2] 枇々木規雄『金融工学と最適化』朝倉書店、2001 年

表 11: 需要の系列とその関連統計量の計算用

i	h_i	d_i	${h_i}^2$	${d_i}^2$	$h_i\, d_i$
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
合計					
平均					

$\Rightarrow$

i	h'_i	d'_i	${h'_i}^2$	${d'_i}^2$	$h'_i\, d'_i$
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
合計					
平均					

表 12: 利潤の系列計算用 （ ＿＿＿＿＿ の場合用）

i	h_i	d_i	$f(h_i)$	$f(d_i)$	$f(h_i)^2$	$f(d_i)^2$	$f(h_i)\,f(d_i)$
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
合計							
平均							

$\Longrightarrow$

i	$f(h'_i)$	$f(d'_i)$	$f(h'_i)^2$	$f(d'_i)^2$	$f(h'_i)\,f(d'_i)$
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
合計					
平均					

ただし、f は（仕入れ値、売値、引き取り値をパラメータとする）利潤の関数。