

統計学 2020 Lecture 2: 確率変数, 確率分布, 期待値, 分散

北門 利英 (東京海洋大学海洋生物資源学科)

2020 年 5 月 20 日

Attention:

- 授業 HP に授業に関する情報をアップデートしていきますので参照ください. (URL: <https://toshihidekitakado.github.io/STAT2020/index.html>)
- 授業に関してわからないことがあれば, メールかリアルタイム接続時 (水曜 13:00-14:00, 質問が多ければ 14:30 まで延長します) に遠慮なく質問してください.

Point:

- 次の用語と意味を理解すること
 - 確率変数とは, どのような値をとるかが確率的にしか分からない変数
 - 確率分布とは, 確率変数の値と確率の関係を規定するメカニズム
 - 期待値は, 確率変数の確率性を考慮した平均値
 - 分散は, 散らばりの大きさの尺度で, 確率変数の (期待値の周りでの) 散らばりの期待値
- 簡単な例を通して, 上記の定義を確認すること
- 平均と分散という言葉について, 観測値の「標本平均と標本分散」と確率変数の「理論平均 (期待値) と理論分散」が異なることを頭に入れておくこと

1 確率変数と確率分布

1.1 はじめに

確率変数と確率分布などを考える上で, 最も簡単な例の一つがサイコロを投げる試行である.

いまサイコロを 1 回振ったときの出る目を考える. もちろんこの変数の値は振るまで分からない. したがって, 出る目は 1~6 の整数を取りうる変数と考えることができる. どのような値をとるか事前には分からない変数であるが, 変数の出方のメカニズム, すなわち取りうる変数のそれぞれ値に対して確率は付与できる. そこでこの変数を Y とおき, その確率を定義すると次のように書ける.

$$P(Y = 1) = P(Y = 2) = \dots = P(Y = 6) = \frac{1}{6}$$

あるいは次のように示してもよい.

$$P(Y = y) = \frac{1}{6} \quad \text{for } y = 1, 2, \dots, 6$$

後で改めて定義するが, このような変数を確率変数とよび, その確率的規則性を確率分布という.

ところで, リアリティーを持たせるために, サイコロを振って出た目の数に応じた金額がもらえるくじがあるとする. すなわち Y 万円もらえる訳である. 一体どれくらいの金額がもらえると期待できるだろうか. このような値のことを期待値と呼ぶが, その計算は次の通り.

$$\begin{aligned} E[Y] &= 1 \times P(Y = 1) + 2 \times P(Y = 2) + \dots + 6 \times P(Y = 6) \\ &= \sum_{y=1}^6 yP(Y = y) = \frac{7}{2} \end{aligned}$$

📖 ここはウォーミングアップです. 用語の説明などは改めて後述します.

1.2 生物の個体数の確率的变化

次のような生物の個体数の確率的变化を考えよう.

例題

ある生物種のメスは寿命が1年で, 子供を産むとすぐに死亡する. 次の世代にメス1個体を産む確率と2個体を産む確率はともに50%と仮定する. このとき, この生物の増加の様子は確率的にしか表現できない. 実験開始時のメスの個体数を1とすると, メスの数は年々どのように変化するであろうか.

高校で習う確率の授業では, 主に事象に対して確率を定義したが, 確率論や統計学の分野では, 観測あるいは数学的な便宜上から, ある変数に対して確率を定義すること多い. 例えば上記の例では, 「あるメス1個体が生計1個体を産む」などの事象に対して確率が付与されていることになる. ここで, $t (= 1, 2, \dots, T)$ 年後のメスの個体数を Y_t と表すことにする. 1年後のメス個体数 Y_1 は1,2のいずれかの値をとりうるが, どの値をとるかは事前には確率的にしかわからない. そこで仮定に基づいて, $Y_1 = 1$ あるいは $Y_1 = 2$ という事象に対して

$$P(Y_1 = 1) = \frac{1}{2}, \quad P(Y_1 = 2) = \frac{1}{2} \tag{1}$$

のように確率を付与する.

2年後のメス個体数は1年目のメス個体数に依存する. すなわち, Y_1 の値が1のとき, 2年後の個体数 Y_2 は1または2の値をとる. 一方, $Y_1 = 2$ のときには Y_2 は2,3,4のいずれかとなる. これを条件付確率を用いて表すと

$$\begin{aligned} P(Y_2 = 1|Y_1 = 1) &= \frac{1}{2}, & P(Y_2 = 2|Y_1 = 1) &= \frac{1}{2} \\ P(Y_2 = 2|Y_1 = 2) &= \left(\frac{1}{2}\right)^2, & P(Y_2 = 3|Y_1 = 2) &= 2 \cdot \left(\frac{1}{2}\right)^2, & P(Y_2 = 4|Y_1 = 2) &= \left(\frac{1}{2}\right)^2 \end{aligned}$$

となる. $Y_1 = 1$ のときには Y_2 は 3, 4 の値をとれないが, これを確率が 0 と考えると, Y_2 の取りうる値と確率の関係は

$$\begin{aligned} P(Y_2 = 1) &= P(Y_2 = 1|Y_1 = 1)P(Y_1 = 1) + P(Y_2 = 1|Y_1 = 2)P(Y_1 = 2) = \frac{2}{8} \\ P(Y_2 = 2) &= P(Y_2 = 2|Y_1 = 1)P(Y_1 = 1) + P(Y_2 = 2|Y_1 = 2)P(Y_1 = 2) = \frac{3}{8} \\ P(Y_2 = 3) &= P(Y_2 = 3|Y_1 = 1)P(Y_1 = 1) + P(Y_2 = 3|Y_1 = 2)P(Y_1 = 2) = \frac{2}{8} \\ P(Y_2 = 4) &= P(Y_2 = 4|Y_1 = 1)P(Y_1 = 1) + P(Y_2 = 4|Y_1 = 2)P(Y_1 = 2) = \frac{1}{8} \end{aligned} \quad (2)$$

で与えられる.

☞ 上の計算を確かめてみて下さい. さらに, ちょっと面倒かもしれませんが, 3 年後についても計算してみるとよい練習になります. 期待を計算して $(3/2)^3$ になれば正解です.

上記の例の変数 Y_1, Y_2 のように, 実験や試行の結果どの事象が生起するかは事前には確率的にしかわからない. そこで, そのような確率的なメカニズムで変化する数を確率変数 (random variable)** とよぶ. 確率変数のとりうるすべての値の集合を標本空間 (sample space) という. この例では $y_1 = \{1, 2\}$ そして $y_2 = \{1, 2, 3, 4\}$ が確率変数 Y_1, Y_2 の標本空間となる. また, この確率変数のとり値に対して確率的規則性を表したものを確率分布 (probability distribution) という.

なお, 確率変数を表す記号は Y に限らず, X, Z など任意の記号を利用してよい. この先の授業のハンドアウトでも, 確率変数を表すのにいろいろな記号を用いることがある.

1.3 モンテカルロシミュレーション

☞ Lecture 2 の際には計算コードは読み飛ばしても構いません.

上記の例において, 3 年目, 4 年目と時間を追って計算することは原理的には可能であるが, 実際に解析的に手で計算をするのは骨が折れる. 10 年目などまで計算しなさい, という宿題が出たとしたら途方に暮れるでしょう.

そのような場合, 乱数を用いたコンピューターによるシミュレーションを利用するのが便利である. この授業では R というソフトを後半で用いる予定であるが, それを利用した例が以下のコードと結果である.

```
T <- 5          # 年の上限設定
Nit <- 10000    # シミュレーションの繰り返し回数
Y <- array(0, c(Nit, T)) # 確率変数の実現値の行列 (繰返回数 × 年数)

for(i in 1:Nit)
{
  # 1 年目のメス個体数
  Y[i,1] <- sample(c(1,2), size=1, prob=c(0.5,0.5))
  # 2 年目以降のメス個体数の計算
  for(t in 2:T)
  {
```

```

# 前年のメス個体毎に子供メス数を割り付けるベクトル
offsprings <- sample(c(1,2), size=Y[i,(t-1)], prob=c(0.5,0.5), replace=T)
# t 年目のメスの個体数
Y[i,t] <- sum(offsprings)
}
}

# シミュレーション 1 回目の結果 Y[1,] を図示
plot(1:T, Y[1,], type="l", col="blue",
     xlab="Year", ylab="#individuals", ylim=c(0,max(Y)))

# シミュレーション 2 回目の結果 Y[i,] を順に図示
for(i in 2:Nit) points(1:T, Y[i,], type="l", col="blue")

# 各年のメス個体数の平均値の計算と図示
MEAN <- apply(Y, 2, mean); MEAN

```

```
[1] 1.4975 2.2437 3.3632 5.0350 7.5745
```

```
points(1:T, MEAN, type="l", lwd=2, col="red")
```

```

# 最終年のメス個体数の確率分布
plot(table(Y[,T])/Nit, xlab="#Female animals", ylab="Probability",
     xlim=c(0,max(Y)), col="orange", main="", lwd=3)

```

例えば、5 年後にメスの個体数が 1 個体のままである確率は厳密には $(1/2)^5 = 1/32$ であるが、モンテカルロ法を用いればそれを近似的に得られることが分かる。また、10 以上である確率も、モンテカルロ法を用いておよそ 30 パーセントと計算できる。

```
mean(Y[,T]==1)
```

```
[1] 0.0291
```

```
mean(Y[,T]>=10)
```

```
[1] 0.2966
```

1.4 確率変数と確率分布

ところで、確率変数には自然数や整数などの離散的な値をとる **離散型確率変数 (discrete random variable)** と数直線上の値など連続な値をとる **連続型確率変数 (continuous random variable)** の 2 種類に大きく分けられる。特に、離散型確率変数に対する確率分布は **離散型確率分布 (discrete probability distribution)** または単に **離散型分布 (discrete distribution)** とよぶ。同様に、連続型確率変数に対する確

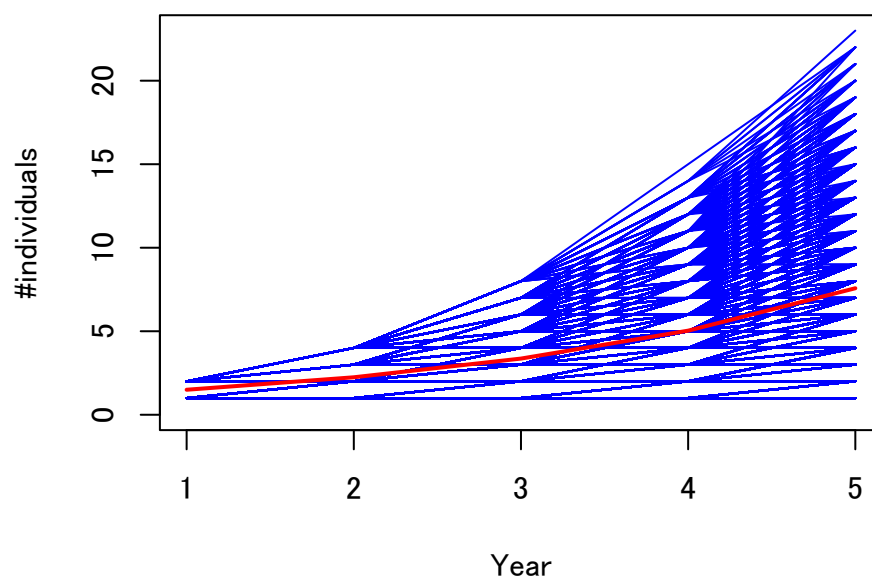


図1 メスの個体数の確率的变化 (T=5 まで)

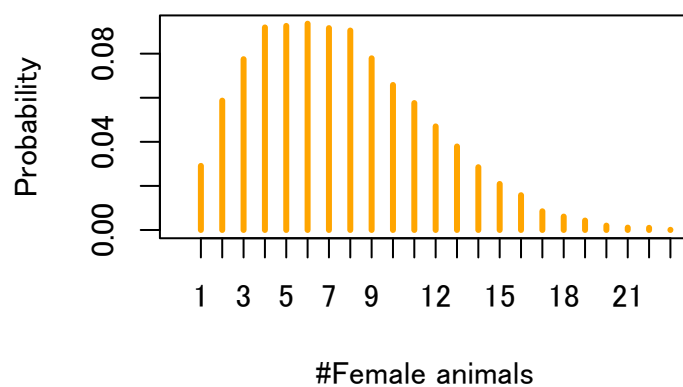


図2 最終年 (T=5) におけるメスの個体数の確率分布の様子 (T=5 まで)

率分布は連続型確率分布 (continuous probability distribution) または単に連続型分布 (continuous distribution) と呼ばれる。

1.4.1 離散型の一番簡単な例

離散型分布 1 [離散型一様分布] 離散型一様分布 (discrete uniform distribution) とは、有限個の要素をもつ標本空間内のすべての値に対して等しい確率を与えた確率分布である。すなわち、標本空間 $\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ をもつ確率変数 X に対して確率関数が

$$f(x) = \frac{1}{n}, \quad x = x_1, x_2, \dots, x_n \quad (3)$$

となる場合をいう。6 面体サイコロでは $x_i = i$ ($i = 1, 2, \dots, 6$) と考えればよい。

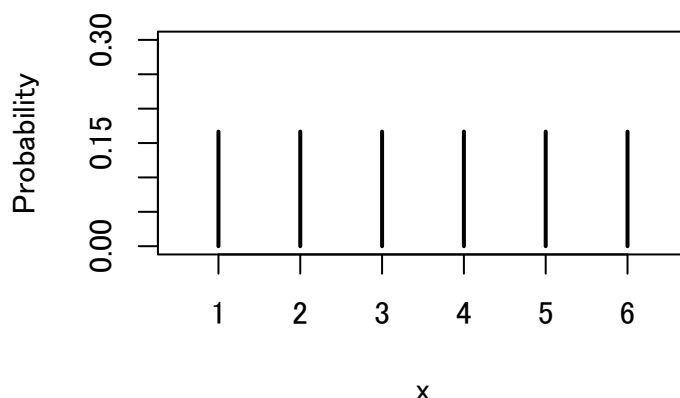


図3 サイコロ 1 回試行の場合の離散一様分布

$f(x)$ のように確率を確率変数の実現値の関数で表す時、これを確率関数 (probability function) とよぶ。この確率関数は離散型の確率分布に対してのみ定義される。

性質 1 [確率関数の性質] 標本空間 $\mathcal{X} = \{x_1, x_2, \dots\}$ をもつ確率変数 X に対して以下が成り立つ。

$$(1) \quad f(x_i) \geq 0$$

$$(2) \quad \sum_{i=1}^{\infty} f(x_i) = 1$$

1.4.2 連続型の一番簡単な例

ところで離散型確率変数の場合には標本空間内のそれぞれの値に対して確率を与えることができた。しかし、連続型の確率変数についてはどうであろうか。例えば、区間 $[0, 1]$ を標本空間とする確率変数 X について考えよう。 X の実現値 $x = 0$ と $x = 1$ に対して、ある小さな確率値 (仮にこれをともに c とする) を割り当てたとする。さらに、この 0 と 1 の間に無数に数があるから、それらに対してもある確率値を割り当てなければな

らない。確率の定義では、すべての全事象の確率は 1 でなければならない。すなわち、標本空間のすべての値に対する確率の和が 1 となる必要がある。しかしながら、どんな小さな確率を割り当てようとも、数直線上には数えられない数の点があり (これを非加算無限という)、したがってその和を 1 で抑えることができないのである。

そこで連続型の場合には分布関数を利用して確率を定義する。確率変数 X がある範囲の値をとる確率を

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

のように定義する。関数 $f(x)$ を X の確率密度関数 (**probability density function, pdf**) または単に密度関数 (**density function**) とよぶ。面積の計算で確率が求まることがわかる。

性質 2 [確率密度関数の性質] 確率変数 X の確率密度関数 $f(x)$ に対して以下が成り立つ。

$$(1) f(x) \geq 0$$

$$(2) \int_{-\infty}^{\infty} f(x)dx = 1$$

連続型分布 1 [連続型一様分布] 連続型一様分布 (**continuous uniform distribution**) または単に一様分布 (**uniform distribution**) とは、標本空間を $X = (\alpha, \beta)$ のどの実現値に対しても確率密度の値が一定である分布をいう。すなわち、確率変数 X が一様分布 $U(\alpha, \beta)$ に従うとは、 X の確率密度関数が

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{if } \alpha < x < \beta \\ 0 & \text{o.w.} \end{cases} \quad (4)$$

となる場合である。連続型分布の場合には 1 点に対する確率が定義されないため、ここでは定義域の両端の値を標本空間には含めないことにする。

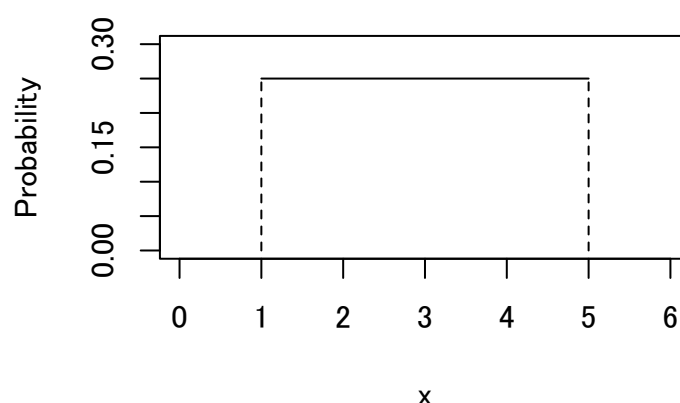


図4 $U(1,5)$ の確率密度関数の概形

2 確率分布の特性値

確率変数あるいは確率分布の性質を表現する指標はいくつかあるが、その中でも最も代表的なのが期待値と分散である。期待値 (expectation) は確率変数のとる平均的な値であり、また分散 (variance) はそのばらつきの程度を表すために用いられる。

2.1 期待値の定義

まずは、期待値の定義を確認しよう。

定義 1 [確率変数の期待値] 確率変数 X の期待値は次のように定義される。

$$E[X] = \begin{cases} \sum_{x=0}^{\infty} x f(x) & (\text{離散型確率分布のとき}) \\ \int_{-\infty}^{\infty} x f(x) dx & (\text{連続型確率分布のとき}) \end{cases} \quad (5)$$

メスの個体数の問題では、1 年目のメス個体数の期待値は

$$E[Y_1] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2},$$

となる。2 年目の期待値も

$$E[Y_2] = 1 \cdot \frac{2}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{2}{8} + 4 \cdot \frac{1}{8} = \frac{9}{4} = \left(\frac{3}{2}\right)^2$$

のように計算できる。なお、 $E[Y_2]$ は条件付期待値

$$\begin{aligned} E[Y_2|Y_1 = 1] &= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2} \\ E[Y_2|Y_1 = 2] &= 2 \cdot \frac{1}{4} + 3 \cdot \frac{2}{4} + 4 \cdot \frac{1}{4} = 3 \end{aligned}$$

を用いて

$$E[Y_2] = E[Y_2|Y_1 = 1] \cdot P(Y_1 = 1) + E[Y_2|Y_1 = 2] \cdot P(Y_1 = 2) = \frac{9}{4} \quad (6)$$

のように計算することもできる。実は、 $E[Y_t] = \left(\frac{3}{2}\right)^t$ であることが示せる。

2.2 モンテカルロシミュレーション (続き, 読み飛ばしてよい)

📖 Lecture 2 の際には計算コードは読み飛ばしても構いません。

今度は 10 年目まで考えると以下の通り。ただし今度はメス 1 個体から、1 個体も生まれない確率が $1/4$, 1 個体生まれるのは $1/2$, そして 2 個体の場合は $1/4$ とする。容易に予想されるように、絶滅することもあり得るし、急激に増加する場合も考えられる。また、個体数の期待値は、1 個体から生まれる期待値が 1 であるから、1 のままである。

```

T <- 10
Nit <- 10000
Y <- array(0, c(Nit, T))
for(i in 1:Nit)
{
  Y[i,1] <- sample(c(0,1,2), size=1, prob=c(0.25, 0.5, 0.25))
  for(t in 2:T)
  {
    offsprings <- sample(c(0,1,2), size=Y[i,t-1], prob=c(0.25, 0.5, 0.25), replace=T)
    Y[i,t] <- sum(offsprings)
  }
}

MEAN <- apply(Y, 2, mean) # 各年における期待値

plot(0,0, type="n", xlab="Year", ylab="Female number",
      xlim=c(1,T), ylim=c(0, 1.1*max(Y)), col="red")
for(i in 1:Nit) points(seq(1,T), Y[i,], type="l", col="darkgray")
points(seq(1,T), MEAN, type="l", lwd=2, col="red")

```

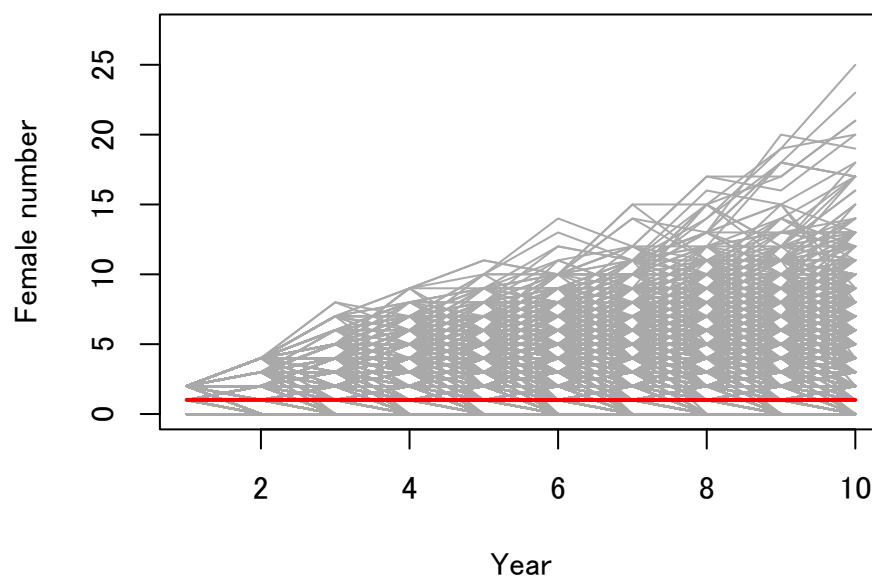


図5 別バージョンのメス個体数の確率的变化 (T=10 まで).

5 年後と 10 年後における個体数のモンテカルロ確率分布をみてみると, 5 年目で約 60 パーセントの確率で, 10 年目には約 75 パーセントの確率で, 個体数が 0 になることが分かる. 期待値は 1 であるが, 大きな個体数にもなる一方で, 0 となる確率も実は大きくなる事で, 期待値のバランスがとられていることが分かる.

```
Probdist <- apply(Y, 2, table)
tmp <- lapply(Probdist, prop.table)
```

```
#5 年後
tmp[[5]]
```

```
      0      1      2      3      4      5      6      7      8      9     10
0.5974 0.1280 0.1183 0.0709 0.0411 0.0218 0.0126 0.0054 0.0025 0.0013 0.0005
11
0.0002
```

```
#10 年後
tmp[[10]]
```

```
      0      1      2      3      4      5      6      7      8      9     10
0.7354 0.0511 0.0577 0.0423 0.0342 0.0223 0.0188 0.0113 0.0085 0.0059 0.0041
11      12      13      14      15      16      17      18      19      20      21
0.0027 0.0014 0.0015 0.0007 0.0003 0.0002 0.0007 0.0002 0.0001 0.0002 0.0002
23      25
0.0001 0.0001
```

2.3 期待値に関する性質

ここで, 期待値の定義について詳細にみていくことにしよう. いま, X のある関数の期待値を考える. この関数は $g(X) = a + bX$ や $h(X) = h_1(X) + \dots + h_n(X)$ など任意のものとする. このとき, これら確率変数の関数に対する期待値はどのように定義されるか. 例えば, Y を離散型とすると,

$$E[g(X)] = E[a + bX] = \sum_{x=0}^{\infty} (a + bx)f(x) = a \sum_{x=0}^{\infty} f(x) + b \sum_{x=0}^{\infty} xf(x) = a + bE[X]$$

となる. 定数の期待値は当然その定数自身であり, 確率変数の定数倍はそのまま定数倍として期待値にも反映される. このように, 期待値には線形性が保たれていることがわかる. また,

$$\begin{aligned} E[h(X)] &= E[h_1(X) + \dots + h_n(X)] = \sum_{x=0}^{\infty} \{h_1(x) + \dots + h_n(x)\}f(x) \\ &= \sum_{x=0}^{\infty} h_1(x)f(x) + \dots + \sum_{x=0}^{\infty} h_n(x)f(x) = E[h_1(X)] + \dots + E[h_n(X)] \end{aligned}$$

となる. このように, 「和の期待値」は「期待値の和」となることが分かる. 以上の性質は連続型分布に関しても同様に成り立つ.

例 1 [St. Petersburg paradox] 公平なコインを投げて初めて表が出るまでの試行数を Y とし, Y の結果に応じて賞金 2^Y (万円) を獲得するゲームを考える. このゲームにおける獲得金額の期待値はいくらであろうか. そこで, まず確率分布を求めてみよう. 初回に表が出る確率が $P(Y = 2) = 1/2$, 2 回目に表が出る確率は $P(Y = 1) = 1/4$ であるから, 一般に

$$P(Y = y) = \frac{1}{2^y}, \quad y = 1, 2, \dots$$

と表せる. すべての確率の和は等比級数の和より

$$\sum_{y=1}^{\infty} P(Y = y) = \sum_{y=1}^{\infty} \left(\frac{1}{2}\right)^y = \frac{1/2}{1 - 1/2} = 1$$

となることが確認される. つぎに, 2^Y の期待値を求めよう. 形式的に期待値の定義にあてはめると,

$$E[2^Y] = \sum_{y=1}^{\infty} 2^y P(Y = y) = \sum_{y=1}^{\infty} \frac{2^y}{2^y} = \infty$$

となり期待値は無限大となる. あなたなら, 参加費がいくらならこのゲームに参加するでしょうか.

2.4 確率変数の分散

定義 2 [確率変数の分散] 確率変数 X の分散 (variance) は次のように定義される.

$$V[X] = \begin{cases} \sum_{x=0}^{\infty} (x - E[X])^2 f(x) & (\text{離散型確率分布のとき}) \\ \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx & (\text{連続型確率分布のとき}) \end{cases} \quad (7)$$

ところで, 分散は言い換えれば「確率変数と期待値の差の 2 乗」の期待値である. そこで, 期待値の性質をとおして改めて分散の定義を見直すと,

$$V[X] = E[(X - E[X])^2] = E[X^2 - 2E[X]X + E[X]^2] \quad (8)$$

であり, 上記の期待値の性質から,

$$V[X] = E[X^2] - 2E[X]E[X] + E[X]^2 = E[X^2] - E[X]^2 \quad (9)$$

が導かれる. 分散は明らかに非負であり, したがって $E[X^2] > E[X]^2$ となる. なお, 分散の平方根を標準偏差 (standard deviation) とよぶ.

☞ なお, $E[X^2] > E[X]^2$ は有名な Schwartz の不等式や Jensen の不等式からも導くことができますが, この授業では深く触れません (テストにもでません).

メス個体数の問題における Y_1 の分散は, 次のように計算できる.

$$V[Y_1] = \left(1 - \frac{3}{2}\right)^2 \cdot \frac{1}{2} + \left(2 - \frac{3}{2}\right)^2 \cdot \frac{1}{2} = \frac{1}{4}.$$

☞ 分散の考え方は, 後の授業で扱う推定や検定で重要な役割を果たしますが, ここでは確率変数の期待値の周りでの散らばりの尺度, とまずは覚えて下さい.

3 観測値の平均と分散

☞ 理論的な平均（期待値）と観測値（データ）の平均とは混同しやすいのですが、徐々に区別できるようになります。今回は、以下を軽く読んでください。

期待値はしばしば理論平均あるいは単に平均値とよばれることもある。これとは別にデータを観測したときの標本平均という言葉も存在する。これらはまったく異なる性質のものである一方で、非常に密接に関連した値でもある。

次にデータの数値的な要約について述べる。いま一般に、ある項目のデータを $x_1, x_2, \dots, x_n$ とするとき、

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (10)$$

をデータの標本平均 (sample mean),

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (11)$$

を標本不偏分散 (sample unbiased variance) または単に 標本分散 (sample variance) という。標本分散の平方根は標本標準偏差 (sample standard deviation) とよばれる。標本分散あるいは標準偏差は平均値の周りでどの程度データがばらつくかを示す尺度であり、データの要約には欠かせない。これらの他に、データを小さい方から順にならべたときに、ちょうど中間の値（偶数のときには中間の 2 つの値の平均値）を **中央値 (median)** あるいはメディアンとよぶ。標本平均も中央値もデータの中心的傾向を測る指標であるが、標本平均が大きな値や小さな値にひきづられやすい一方で、中央値はその影響をあまり受けない。中央値は 50% 点ともよばれるが、同様に 25% 点, 75% 点もデータの散らばりを示すのにしばしば利用される。

サイコロを n 回投げるとしよう。第 i 回目 ($i = 1, 2, \dots, n$) の試行の結果を表す確率変数を X_i とする。 X_i の確率分布は離散型一様分布

$$P(X_i = k) = \frac{1}{6} \quad (k = 1, 2, \dots, 6; i = 1, 2, \dots, n)$$

となる。確率変数 X_i の期待値はどの i に対しても“定数” 3.5 である。ところで、標本平均とは

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (12)$$

であり、これは確率変数のとる値によって変化する値である。したがって、定数である期待値とは対照的に、標本平均それ自身が確率変数であり、必然的に確率分布をもつ。 n の値が小さいとき（例えば $n = 2, 3$ などのとき）には $\bar{X}$ の確率分布を求めることは容易であるが、 n が大きい時には、標本空間自体が大きくなり（とり得る値がどんどん増え）、したがって確率分布それ自身を詳細に求めることにあまり意味はないであろう。後に述べる中心極限定理の利用が実用的である。ところが、標本平均の期待値は簡単に求められる。なぜならば、

$$E[\bar{X}] = E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n E[X_i] = E[X_1]$$

であり、「同一な確率分布から得られる標本平均の期待値は、元の確率変数の期待値と一致する」ことが分かる。