統計の分析と利用

2. 確率変数と確率分布
確率変数の期待値(平均)と分散

堀田 敬介

2013/11/1,Fri.～

試行とは？

- 試行
  - 何かの行為により「偶然による」ひとつの結果を導き出す

〔例〕

さいころ投げ

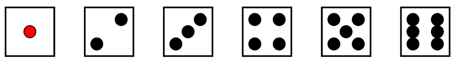
コイン投げ

〔例〕身長の測定，じゃんけん，宝くじを買う，アンケート調査，製品品質検査，etc.

確率変数とは？

試行してみないと何が出るかはわからない！
とりうる値はわかっている

- 確率変数　random variable
  - それがとる各値に対し確率が与えられている変数
    - 例：さいころ投げ



試行結果

$X=$	1	2	3	4	5	6	確率変数の値
$P(X=x_k)=$	1/6	1/6	1/6	1/6	1/6	1/6	確率

Probability(確率)の頭文字 P

ex) $P(X=3)=1/6$
「Xが3となる確率が1/6」と読む

確率変数とは？

- 例：コイン投げ

試行結果

$X=$	表	裏	確率変数の値
$P(X=x_k)=1/2$		1/2	確率

ex) $P(X=\text{裏})=1/2$
「Xが裏となる確率が1/2」と読む

- 一般に，確率変数の確率は以下のように表現される

$$P(X=x_k)=p_k \quad (k=1,2,\cdots)$$

ただし，$p_k \ge 0\,(k=1,2,\cdots)$，$\sum_{k=1}^{\infty} p_k = 1$ である．

確率はすべて0以上　　全ての確率を足すと1

演習1

- 確率変数
  - 2枚の硬貨を投げ，2枚の「裏表」の組合せ（4通り）を考える．この確率変数 X のとる値と，その値が出る確率を求めよ．ただし，どちらの硬貨も表裏同確率で出るとする
    - 例）1枚目が「表」で，2枚目も「表」の時

✓ $P(X=$表・表$) = 1/4$ （$= 1/2 \times 1/2$）
✓ $P(X=$表・裏$) = 1/4$ （$= 1/2 \times 1/2$）
✓ $P(X=$裏・表$) = 1/4$ （$= 1/2 \times 1/2$）
✓ $P(X=$裏・裏$) = 1/4$ （$= 1/2 \times 1/2$）

- 表の形にまとめると…

X	表・表	表・裏	裏・表	裏・裏
$P(X)$	1/4	1/4	1/4	1/4

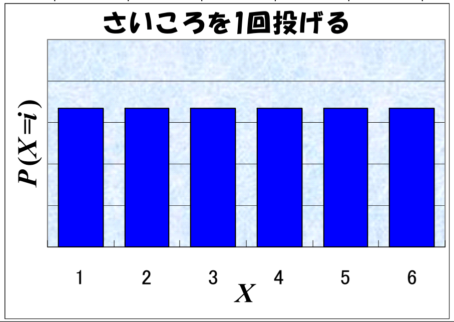
← 確率変数 X のとりうる全ての値
← その確率

これを確率分布という

確率分布 probability distribution

- 確率分布 probability distribution
  - 例：さいころを1回投げる

X	1	2	3	4	5	6
$P(X)$	1/6	1/6	1/6	1/6	1/6	1/6

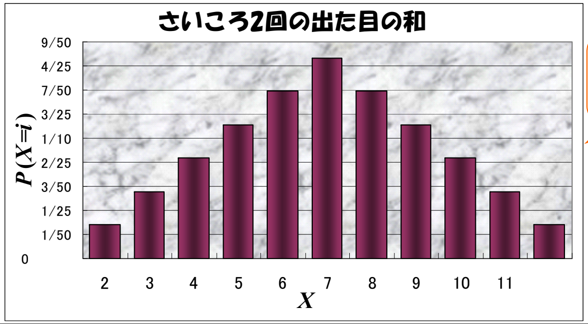


一様分布

確率分布 probability distribution

- 確率分布
  - 例：さいころを2回投げ，出た目の和

X	2	3	4	5	6	7	8	9	10	11	12
$P(X)$	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36



二項分布

参考：確率分布 probability distribution

- 離散（型）確率分布 discrete distribution
  - 可算集合 $\{x_1, x_2, \ldots\}$ の中の値を取る確率変数 X は離散型 discrete type といわれる．このとき，それぞれの値の確率

$$f(x_k) := P(X = x_k) \quad (k = 1, 2, \cdots)$$

を X の確率分布 probability distribution という．

ただし，
$$\begin{cases} f(x_k) \geq 0 \quad (k = 1, 2, \cdots), \\ \sum_{k=1}^{\infty} f(x_k) = 1 \end{cases}$$

確率分布
probability distribution

一般的な定義

確率変数の期待値・分散

- **期待値 expectation, expected value**
  - 確率変数 X の期待値
    - 例：コインを3回投げて表が出る回数の確率分布

X	0	1	2	3
$P(X)$	1/8	3/8	3/8	1/8

コインを3回投げると，平均して1.5回表が出ることが期待される

    - 期待値

$$E(X) = (0\times\frac{1}{8}) + (1\times\frac{3}{8}) + (2\times\frac{3}{8}) + (3\times\frac{1}{8}) = \frac{3}{2}$$

$$\left(\begin{aligned} E(X) &= (\frac{0\times1}{8}) + (\frac{1\times3}{8}) + (\frac{2\times3}{8}) + (\frac{3\times1}{8}) \\ &= \frac{0}{8} + \frac{1+1+1}{8} + \frac{2+2+2}{8} + \frac{3}{8} \\ &= \frac{0+1+1+1+2+2+2+3}{8} \end{aligned} \right)$$

  - 確率変数Xの期待値

$$E(X) = \sum_x x \cdot f(x)$$

期待値は算術平均を計算しているのと同じ

確率変数の期待値・分散

- **分散 variance**
  - 確率変数 X の分散

分散（ばらつき）＝平均（期待値）からのずれの平均

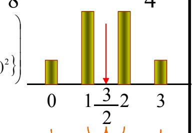
$$V(X) = E(\{X - E(X)\}^2) \qquad \left(V(X) = E(X^2) - E(X)^2\right)$$

    - 例：コインを3回投げて表が出る回数の分布の分散は?

$$V(X) = \frac{1}{8}\cdot(0-1.5)^2 + \frac{3}{8}\cdot(1-1.5)^2 + \frac{3}{8}\cdot(2-1.5)^2 + \frac{1}{8}\cdot(3-1.5)^2 = \frac{3}{4}$$

$$\left(\begin{aligned} V(X) &= \frac{1}{8}\cdot\{1\times(0-1.5)^2 + 3\times(1-1.5)^2 + 3\times(2-1.5)^2 + 1\times(3-1.5)^2\} \\ &= \frac{1}{8}\cdot\{(0-1.5)^2 + (1-1.5)^2 + (1-1.5)^2 + (1-1.5)^2 + (2-1.5)^2 + (2-1.5)^2 + (2-1.5)^2 + (3-1.5)^2\} \end{aligned} \right)$$

通常の分散を計算しているのと同じ



平均的にどの程度散らばっているか？

  - 確率変数Xの分散

$$V(X) = \sum_x \left(x - E(X)\right)^2 f(x)$$

確率変数の期待値・分散

- **標準偏差 standard deviation**
  - 確率変数Xの標準偏差

標準偏差＝分散の平方根

$$D(X) = \sqrt{V(X)}$$

    - 例：コインを3回投げて表が出る回数の分布の標準偏差は?

X	0	1	2	3
$P(X)$	1/8	3/8	3/8	1/8

$$D(X) = \sqrt{V(X)} = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2}$$

補足：確率変数の歪度・尖度

- **歪度 skewness**
  - 確率変数Xの確率分布の非対称性の指標
  - 歪度＝α_3. ただし，

$$\alpha_3 = E(X-\mu)^3 / \sigma^3$$

$$\left(\mu := E(X),\ \sigma^2 := V(X)\right)$$

歪度の数値の意味

$$\begin{cases} \alpha_3 > 0 \ldots \text{右の裾が長い} \\ \alpha_3 < 0 \ldots \text{左の裾が長い} \\ |\alpha_3| \quad \ldots \text{歪みの程度} \end{cases}$$

- **尖度 kurtosis**（超過係数 coefficient of excess）
  - 確率変数Xの確率分布の尖り具合を表す指標
  - 尖度＝α_4―3. ただし，

$$\alpha_4 = E(X-\mu)^4 / \sigma^4$$

正規分布がα_4=3なので，これと比較

尖度の数値の意味

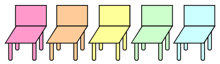
$$\begin{cases} \alpha_4 - 3 > 0 & \text{正規分布より尖っている} \\ \alpha_4 - 3 < 0 & \text{正規分布より丸く鈍い形} \end{cases}$$

Coffee Break!

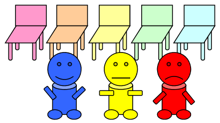
階乗・順列・組合せ・重複組合せ

factorial permutation combination repeated combination

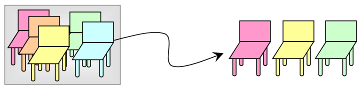
- 異なる5脚の椅子を一列に並べる方法は何通り?



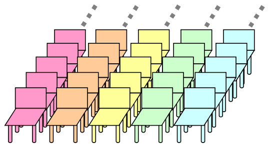
- 5脚の椅子に3人の学生が座る座り方は何通り?



- 5脚の椅子から使いたい3脚を選ぶ方法は何通り?

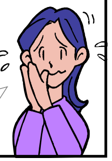


- 異なる5種類の椅子から8脚選ぶ選び方は何通り?



- 異なる5種類の椅子から少なくとも1種類ずつ8脚選ぶ選び方は?

それぞれどんな計算になるのかしら?



Coffee Break!

階乗 factorial

$5! = 5\cdot4\cdot3\cdot2\cdot1 = 120$

※）Excel関数：FACT(5)

順列 permutation

${}_5P_3 = 5\cdot4\cdot3 = 60$

※）Excel関数：PERMUT(5,3)

組合せ combination

${}_5C_3 = \dfrac{5\cdot4\cdot3}{3\cdot2\cdot1} = 10$

※）Excel関数：COMBIN(5,3)

重複組合せ repeated combination

$${}_5H_8 = {}_{8+5-1}C_8 = \frac{12\cdot11\cdot10\cdot9}{4\cdot3\cdot2\cdot1} = 495$$

選ぶ椅子8脚と5種類の椅子の間に置く仕切り(5−1)=4を用意，8+5-1から8脚の椅子を置く場所を選ぶ組合せ

$${}_5H_3 = {}_{3+5-1}C_3 = \frac{7\cdot6\cdot5}{3\cdot2\cdot1} = 35$$

5脚は決まるから，残り3脚を上記と同じ方法で選べば良い

その他に…

円順列 circular permutation

- 異なる5脚の椅子を円形に並べる方法は何通り? $\dfrac{5!}{5} = (5-1)! = 24$

1つずらして同じものが5個ある

数珠順列 necklace permutation

- 異なる5脚の椅子を宇宙空間に円形に並べる方法は何通り? $\dfrac{(5-1)!}{2} = 12$

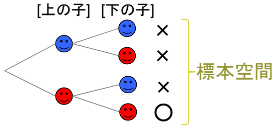
円順列から裏表（逆回り）を除く

Coffee Break!

ベイズ推定初歩

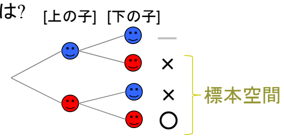
- 2人の子供を持つ家庭がある（注：男女の生まれる確率は各々1/2とする）
- *Q1*：2人とも女の子である確率は?

$\dfrac{1}{4}$



- *Q2*：1人が女の子の時，2人とも女の子である確率は?

$\dfrac{1}{3}$



- *Q3*：1人が女の子で名前がAliceの時，2人とも女の子である確率は?

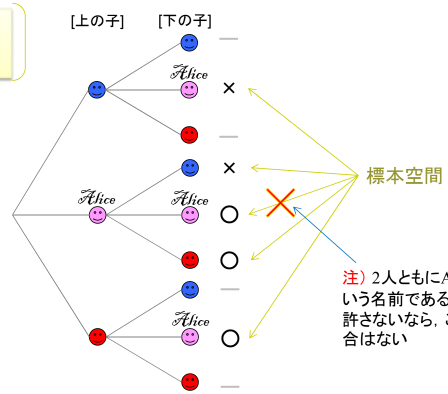
$\dfrac{3}{5}$ or $\dfrac{1}{2}$

Coffee Break!

ベイズ推定初歩

- *Q3*：1人が女の子で名前がAliceの時，2人とも女の子である確率は?

$\dfrac{3}{5}$ or $\dfrac{1}{2}$



注）女の子がAliceと名付けられる場合とそうでない場合の確率を同じとみなしている

実際にはAliceと名付けられる確率は，そうでない場合よりずっと低いだろう

また，ここでは，どちらかの名前がAliceであると判明した（新しい情報が付加された）場合の話であることに注意

確率分布
probability distribution

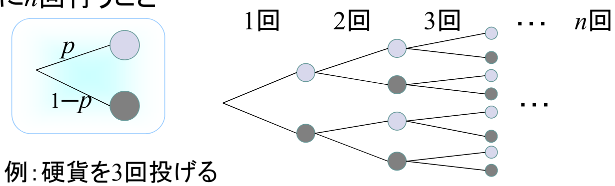
離散(型)分布 discrete distribution

★二項分布 binomial distribution
★ポアソン分布 Poisson distribution

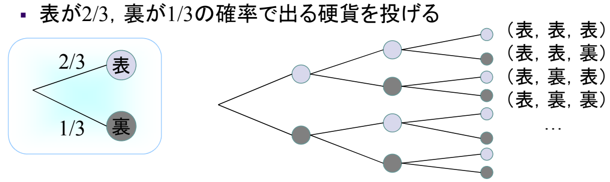
★(離散)一様分布 uniform distribution
★幾何分布 geometric distribution
★負の二項分布 negative binomial distribution
★超幾何分布 hypergeometric distribution

離散型分布 discrete distribution

- ベルヌーイ試行
  - 2通りの結果しかない観測があり，これを同条件で独立にn回行うこと

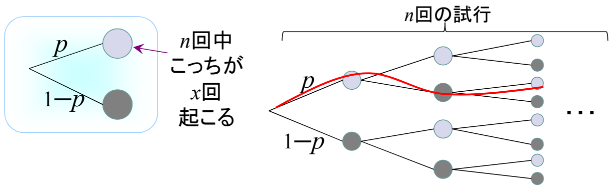


  - 例：硬貨を3回投げる
    - 表が2/3，裏が1/3の確率で出る硬貨を投げる



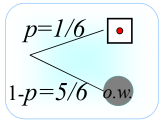
1. 二項分布

- **二項分布 binomial distribution** $Bi(n, p)$
  - ベルヌーイ試行で，片方の結果が起こる回数の確率
  - 確率pをもつ事象がn回の施行中x回起こる確率



  - 例1） 3割打者(p=3/10)が打席に423回立つ(n=423)．このとき，ヒットを打つ回数(x回)の確率分布が二項分布 $Bi(423, 0.3)$
  - 例2） ゴール決定率20%のストライカー(p=0.2)が，175本シュートを放つ(n=175)．このとき，シュートが決まる回数(x回)の確率分布が二項分布 $Bi(175, 0.2)$

1. 二項分布



- **二項分布 binomial distribution**
  - 例：サイコロを3回投げて1の目がx回出る確率は…

1箇所の1の置き方
$_3C_1$(=3通り)
1が1回出る
$\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^2$

2箇所の1の置き方
$_3C_2$(=3通り)
1が2回出る
$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1$

0回出る確率 … $(5/6)^3 = {}_3C_0\,(1/6)^0\,(5/6)^3 = 125/216$
1回出る確率 … ${}_3C_1\,(1/6)^1\,(5/6)^2 = {}_3C_1\,(1/6)^1\,(5/6)^2 = 75/216$
2回出る確率 … ${}_3C_2\,(1/6)^2\,(5/6)^1 = {}_3C_2\,(1/6)^2\,(5/6)^1 = 15/216$
3回出る確率 … $(1/6)^3 = {}_3C_3\,(1/6)^3\,(5/6)^0 = 1/216$

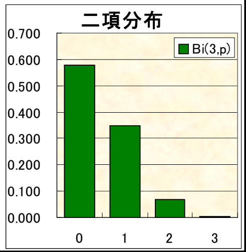
【注）${}_3C_3 = {}_3C_0 = 1$】

二項分布$Bi(3, 1/6)$の確率を計算する式

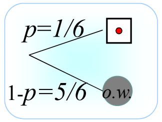
$$f(x) = {}_3C_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{3-x} \quad (x = 0,1,2,3)$$

二項分布$Bi(3, 1/6)$の確率分布

X	0	1	2	3
$P(X)$	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$



1. 二項分布



- **二項分布 binomial distribution**
  - 例：サイコロを3回投げて1の目がx回出る確率は…

1箇所の1の置き方
└ $_3C_1$(=3通り)
1が1回出る
└ $\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^2$

2箇所の1の置き方
└ $_3C_2$(=3通り)
1が1回出る
└ $\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1$

0回出る確率 … $(5/6)^3 = {}_3C_0\,(1/6)^0\,(5/6)^3 = 125/216$
1回出る確率 … ${}_3C_1\,(1/6)^1\,(5/6)^2 = {}_3C_1\,(1/6)^1\,(5/6)^2 = 75/216$
2回出る確率 … ${}_3C_2\,(1/6)^2\,(5/6)^1 = {}_3C_2\,(1/6)^2\,(5/6)^1 = 15/216$
3回出る確率 … $(1/6)^3 = {}_3C_3\,(1/6)^3\,(5/6)^0 = 1/216$

【注】$_3C_3 = {}_3C_0 = 1$】

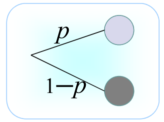
二項分布$Bi(3, 1/6)$の確率を計算する式

$$f(x) = {}_3C_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{3-x} \quad (x = 0,1,2,3)$$

一般的に，二項分布$Bi(n, p)$の確率を計算する式
（n回のベルヌーイ試行を行い，確率pの事象がx回起こる）

$$f(x) = {}_nC_x p^x (1-p)^{n-x} \quad (x = 0,1,2,\cdots,n)$$

1. 二項分布



- **二項分布 binomial distribution**
  - 例1） 3割打者（p=3/10）が打席に423回立つ（n=423）．このとき，ヒットを打つ回数（x回，確率変数X）の確率分布が二項分布 $Bi(423, 0.3)$

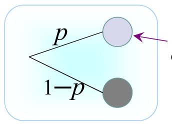
$f(x)=$？
$$f(x) = {}_{423}C_x\left(\frac{3}{10}\right)^x\left(\frac{7}{10}\right)^{423-x} \quad (x = 0,1,2,\cdots,423)$$
50本ヒット確率？ $P(X=50) = f(50) = {}_{423}C_{50}(3/10)^{50}(7/10)^{373}$
200本以上ヒット確率？ $P(X \geq 200) = \sum_{x=200}^{423} f(x) = f(200) + f(201) + \cdots + f(423)$

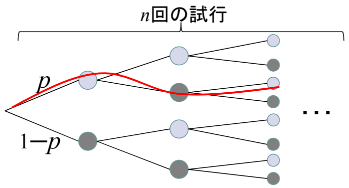
  - 例2） ゴール決定率20%のストライカー（p=0.2）が，175本シュートを放つ（n=175）．このとき，シュートが決まる回数（x回，確率変数X）の確率分布が二項分布 $Bi(175, 0.2)$

$f(x)=$？
$$f(x) = {}_{175}C_x\left(\frac{2}{10}\right)^x\left(\frac{8}{10}\right)^{175-x} \quad (x = 0,1,2,\cdots,175)$$
15本以上30本以下ゴール確率？
$$P(15 \leq X \leq 30) = \sum_{x=15}^{30} f(x) = f(15) + f(16) + \cdots + f(30)$$

1. 二項分布

- **二項分布 binomial distribution** $Bi(n, p)$
  - ベルヌーイ試行で，片方の結果が起こる回数の確率
  - 確率pをもつ事象がn回の施行中x回起こる確率





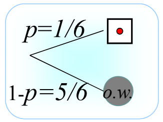
  - 確率分布

$$f(x) = {}_nC_x p^x (1-p)^{n-x} \quad (x = 0,\cdots,n)$$

  - 期待値・分散

$$\begin{cases} E(X) = np, \\ V(X) = np(1-p) \end{cases}$$

1. 二項分布



- **二項分布 binomial distribution**
  - 例：サイコロを3回投げて1の目がx回出る確率は…

3箇所に1個1を置く
└ $_3C_1$
1が1回出る
└ $\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^2$

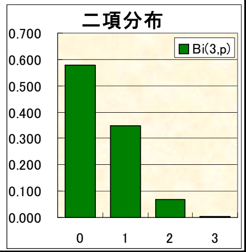
0回出る確率 … $(5/6)^3 = {}_3C_0\,(1/6)^0\,(5/6)^3 = 125/216$
1回出る確率 … ${}_3C_1\,(1/6)^1\,(5/6)^2 = {}_3C_1\,(1/6)^1\,(5/6)^2 = 75/216$
2回出る確率 … ${}_3C_2\,(1/6)^2\,(5/6)^1 = {}_3C_2\,(1/6)^2\,(5/6)^1 = 15/216$
3回出る確率 … $(1/6)^3 = {}_3C_3\,(1/6)^3\,(5/6)^0 = 1/216$

$$f(x) = {}_3C_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{3-x} \quad (x = 0,1,2,3)$$

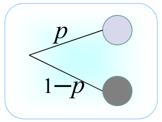
$$\begin{cases} E(X) = 3\cdot\frac{1}{6} = \frac{1}{2} \\ V(X) = 3\cdot\frac{1}{6}\cdot\frac{5}{6} = \frac{5}{12} \end{cases}$$

確率分布

X	0	1	2	3
$P(X)$	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$



1. 二項分布

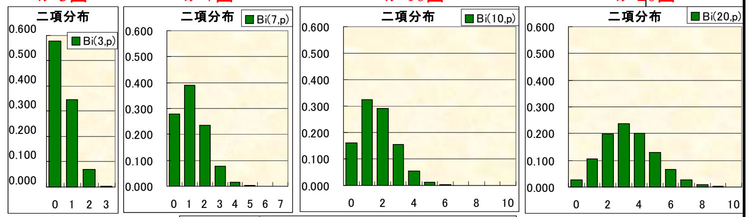


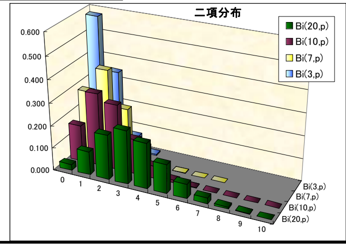
- **二項分布 binomial distribution**
  - 例1） 3割打者（p=3/10）が打席に423回立つ（n=423）．このとき，ヒットを打つ回数（x回）の確率分布が二項分布 $Bi(423, 0.3)$
    $f(x) = {}_{423}C_x \left(\frac{3}{10}\right)^x \left(\frac{7}{10}\right)^{423-x} \quad (x = 0,1,2,\cdots,423)$
    期待値：$E(X)$＝?　$E(X) = np = 423 \times (3/10) = 126.9$
    分散：$V(X)$＝?　$V(X) = np(1-p) = 126.9 \times (7/10) = 88.83$
  - 例2） ゴール決定率20%のストライカー（p=0.2）が，175本シュートを放つ（n=175）．このとき，シュートが決まる回数（x回）の確率分布が二項分布 $Bi(175, 0.2)$
    $f(x) = {}_{175}C_x \left(\frac{2}{10}\right)^x \left(\frac{8}{10}\right)^{175-x} \quad (x = 0,1,2,\cdots,175)$
    期待値：$E(X)$＝?
    標準偏差：$D(X)$＝?
    $E(X) = np = 175 \times (2/10) = 35$
    $D(X) = \sqrt{V(X)} = \sqrt{np(1-p)} = \sqrt{35 \times (8/10)} = 4\sqrt{7}$

1. 二項分布

- pが一定（p=1/6）で，試行回数を増やす（nが増える）と…

n=3回　n=7回　n=10回　n=20回



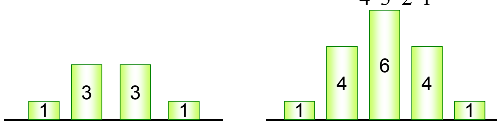


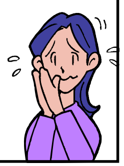
パスカルの三角形と二項係数

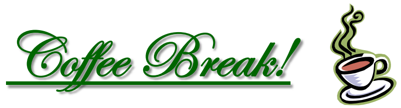
$${}_3C_0 = 1 \qquad {}_3C_1 = \frac{3}{1} = 3 \qquad {}_3C_2 = \frac{3\cdot 2}{2\cdot 1} = 3 \qquad {}_3C_3 = \frac{3\cdot 2\cdot 1}{3\cdot 2\cdot 1} = 1$$

$${}_4C_0 = 1 \qquad {}_4C_1 = \frac{4}{1} = 4 \qquad {}_4C_2 = \frac{4\cdot 3}{2\cdot 1} = 6 \qquad {}_4C_3 = \frac{4\cdot 3\cdot 2}{3\cdot 2\cdot 1} = 4 \qquad {}_4C_4 = \frac{4\cdot 3\cdot 2\cdot 1}{4\cdot 3\cdot 2\cdot 1} = 1$$

…







1
1 1 ⟷ ${}_1C_0 \quad {}_1C_1$
1 2 1 ⟷ ${}_2C_0 \quad {}_2C_1 \quad {}_2C_2$
1 3 3 1 ⟷ ${}_3C_0 \quad {}_3C_1 \quad {}_3C_2 \quad {}_3C_3$
1 4 6 4 1 ⟷ ${}_4C_0 \quad {}_4C_1 \quad {}_4C_2 \quad {}_4C_3 \quad {}_4C_4$
1 5 10 10 5 1 ⟷ ${}_5C_0 \quad {}_5C_1 \quad {}_5C_2 \quad {}_5C_3 \quad {}_5C_4 \quad {}_5C_5$
⋮　⋮

☆ 二項定理，二項係数
$(a+b)^n$ の各項の係数

$$\begin{cases} (a+b)^2 = a^2 + 2ab + b^2 \\ (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 \\ (a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \\ (a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5 \\ \vdots \end{cases}$$

☆ 組合せ数の和法則
$${}_nC_k = {}_{n-1}C_{k-1} + {}_{n-1}C_k \quad (\text{for } 1 \le k \le n-1)$$

☆ $${}_nC_k = \sum_{n-k-1 \le m \le n-1} {}_mC_{m+k+1-n} = \sum_{k-1 \le m \le n-1} {}_mC_{k-1}$$

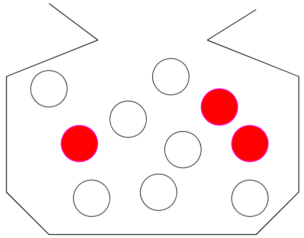
10＝1＋3＋6
10＝1＋2＋3＋4

1. 二項分布

- 二項分布の例
  - 例1：製品ラインの不良品抜き取り検査
    - 不良率p=0.5%のロットから独立に1個ずつランダム抜取り検査をした時に検出される不良品数x の従う分布
    - 参考）不良率の期待値 $E(x/n)=np/n=p$
  - 例2：袋から球を取り出す
    - 赤玉3，白玉7入っている袋から1つ取り出しては戻すという行為をn回行ったとき，赤玉が5回出る確率は?
  - 例3：サイコロをn回投げて偶数の目が出る回数の従う分布
    - サイコロを5回投げて偶数が出る回数の確率分布を求めよ

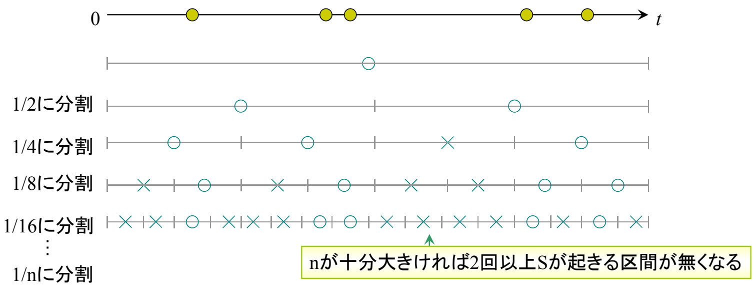
演習3

- 二項分布を求めよう
  - 赤玉3個，白玉7個入っている袋から1つ取り出しては戻すという行為を5回行ったとき，赤球が出る回数の確率分布（二項分布）を求めてみよう！



2. ポアソン分布

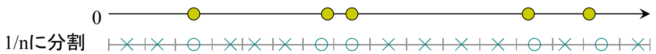
- ポアソン分布 Poisson distribution
  - ある時間帯の中で，ある事象が何回起きるか？
    - 例：電話番号案内
      - 事象S＝「通話がある」



⇒ P(ある区間でSが2回以上起きる確率)=0とする

2. ポアソン分布

- ポアソン分布 Poisson distribution



$\left\{\begin{array}{l}\text{各区間でSが起きる確率} \quad : p \\ \text{各区間でSが起きない確率} : q=1-p\end{array}\right\}$ とする ← ベルヌーイ試行とみなす

⇒ Sが起きる回数は二項分布 $Bi\ (n, p)$ に従う ← 確率pの事象がn回の試行の中でS回起こる

ところで，この時間内にSが起きる回数の期待値をλとおくと…

$\lambda = np$ （np=二項分布の期待値$E(X)$ を 一定に保つ）

よって，S が k回 起きる確率は，

$$
\begin{aligned}
&{}_nC_k p^k (1-p)^{n-k} \\
&= {}_nC_k \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!}\left((1-\frac{1}{n})\cdots\cdots(1-\frac{k-1}{n})\right)\left(1-\frac{\lambda}{n}\right)^{-k}\left(1-\frac{\lambda}{n}\right)^n \xrightarrow[n\to\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}
\end{aligned}
$$

$$\left(\because \lim_{n\to\infty}(1-\frac{\lambda}{n})^n = e^{-\lambda}\right)$$

2. ポアソン分布

- ポアソン分布 $P_o(\lambda)$
  - 確率分布

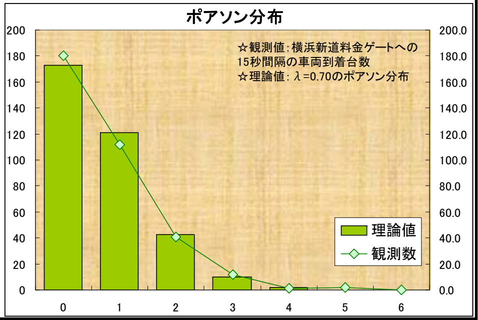
$$f(x)=e^{-\lambda}\frac{\lambda^x}{x!},\ (x=0,1,2,\cdots)$$

  - 期待値・分散

$$\begin{cases} E(X)=\lambda, \\ V(X)=\lambda \end{cases}$$

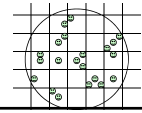
例では…

$$\begin{cases} E(X)=0.70, \\ V(X)=0.70 \end{cases}$$



2. ポアソン分布

- ポアソン分布の例
  - 例1：飛行機事故（事故はめったに起きない）
    - 飛行機事故の確率1/10万，飛行機搭乗回数を1万回としたとき，一度も事故にあわない確率は?
  - 例2：大量生産品の不良品数（めったにない）
    - 不良率が1/10000の生産ラインで1万個生産したとき不良品が3個以上出る確率は?
  - 例3：爆撃命中数（めったに当たらない）
    - 第二次大戦中のドイツ軍の砲弾命中精度はλ=0.93のポアソン分布に従うという．1000発打って1発当たる確率は?
  - 例4：薬の副作用
    - 副作用の確率が1/200の薬を5000人が服用したとき，30人以上に副作用が出る確率は?
  - 例5：生物・植物の生態・繁茂状況を示す分布
    - 単位面積あたりのバクテリアの個数



2. ポアソン分布

- 二項分布 と ポアソン分布
  - 二項分布においてある事象が起こる確率が非常に小さい場合に適用できる分布
    - 例：工場の生産ラインでの不良率が1/500のとき，1000個の製品を作ったときx個不良品だった

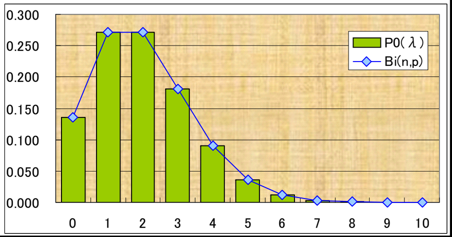
$$f(x)={}_{1000}C_x\left(\frac{1}{500}\right)^x\left(1-\frac{1}{500}\right)^{1000-x}$$

二項分布

ポアソン分布

$$f(x)=e^{-2}\frac{2^x}{x!}$$

1000個のうち，平均的に2個不良品



2. ポアソン分布

- 二項分布 から ポアソン分布 へ

ポアソンの小数の法則
Poisson's law of small numbers

$np=\lambda$

$${}_nC_x p^x(1-p)^{n-x}\to e^{-\lambda}\frac{\lambda^x}{x!}\quad \begin{pmatrix} n\to\infty, \\ p\to 0 \end{pmatrix}$$

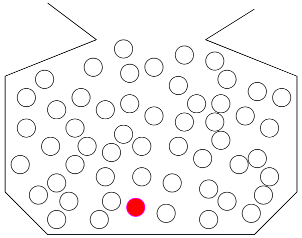
二項分布　ポアソン分布

- 例：ある工場の生産ラインでの不良率が1/500である．1000個の製品を作ったときx個不良品だった．

二項分布	ポアソン分布
$f(0)=0.135065,$	$f(0)=0.135335,$
$f(1)=0.270670,$	$f(1)=0.270671,$
$f(2)=0.270942$	$f(2)=0.270671,$
$f(3)=0.180628,$	$f(3)=0.180447,$
$f(4)=0.090223,$	$f(4)=0.090224,$
$\vdots$	$\vdots$

演習4

- ポアソン分布を求めよう
  - 赤玉1個，白玉99個入っている袋から1つ取り出しては戻すという行為を5回行ったとき，赤球が出る回数の確率分布（ポアソン分布）を求めてみよう！



その他の離散型分布

- (離散)一様分布 uniform distribution (of discrete type)
  - すべての確率が等しい分布
    - 例：さいころを1回投げる

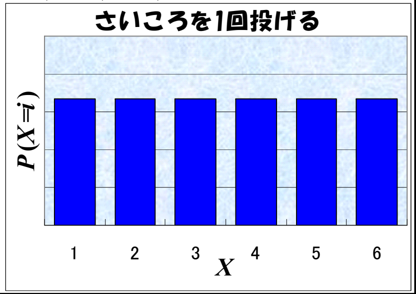
X	1	2	3	4	5	6
$P(X)$	1/6	1/6	1/6	1/6	1/6	1/6

  - 確率分布

$$f(x)=\frac{1}{n},\ (x=1,\cdots,n)$$

  - 期待値・分散

$$\begin{cases} E(X)=\dfrac{n+1}{2}, \\ V(X)=\dfrac{(n^2-1)}{12} \end{cases}$$

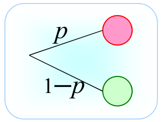


その他の離散型分布

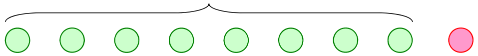
- 幾何分布 geometric distribution
  - ベルヌーイ試行において，試行回数を決めずに初めてある事象が起こるまでの試行回数を x とするときの X=x の確率分布

$$f(x)=p(1-p)^{x-1}\quad (x=1,2,\cdots)$$

幾何数列（等比数列）の形なので，幾何分布とよばれる



x-1回　　x 回目

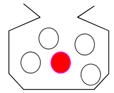


  - 幾何分布は，時間を離散的に(1,2,3,…)考えるとき，初めて何かが起こるまで待つ時間の長さの確率分布〔（離散的な）待ち時間分布〕

その他の離散型分布

- 幾何分布
  - 確率分布

$$f(x)=p(1-p)^{x-1},\ (x=1,2,\cdots)$$

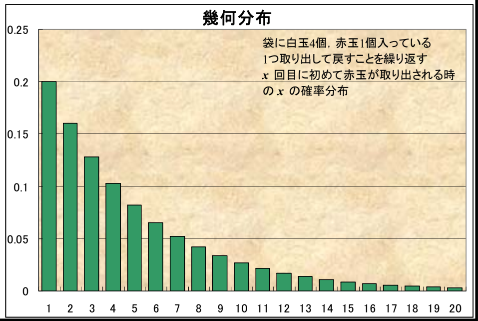


  - 期待値・分散

$$\begin{cases} E(X)=\dfrac{1}{p}, \\ V(X)=\dfrac{1-p}{p^2} \end{cases}$$

例では…

$$\begin{cases} E(X)=\dfrac{1}{1/5}=5.00, \\ V(X)=\dfrac{(1-1/5)}{(1/5)^2}=20.00 \end{cases}$$

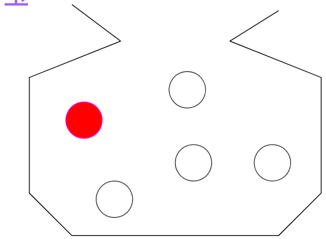


その他の離散型分布

- 幾何分布の例
  - 例1：災害の到来
    - ある1年に風水害が起こる確率が1/25であるとする．風水害が起こるのは平均何年に1回か?
    - 上記と同じ災害が20年以内に起こる確率は?
  - 例2：袋から…
    - 白玉4つ，赤玉1つが入っている袋がある．1つ取り出して元に戻すという試行を繰り返したとき，10回目に初めて赤玉が取り出される確率は?
  - 例3：ドアを開けられる鍵を見つけよう！
    - n個の鍵束を持っている．かぎ束からひとつ鍵を取り出しドアを開けるとき，何回目で開くか?　ただし，試した鍵は1回毎に鍵束に戻すこととする

演習5

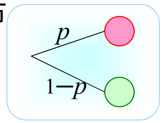
- 幾何分布を求めよう
  - 白玉4つ，赤玉1つが入っている袋がある．1つ取り出して元に戻すという試行を繰り返したとき，初めて赤玉が取り出される回の確率分布（幾何分布）を求めてみよう！　10回目に初めて赤玉が取り出される確率はどれだけか？



その他の離散型分布

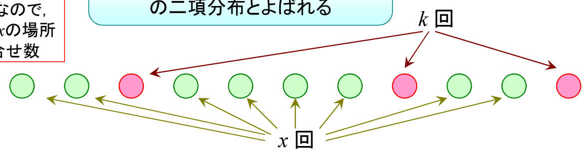
- 負の二項分布 negative binomial distribution
  - ある事象がk回起こるまでのもうひとつの事象の回数をxとしたときのX=x (x=0,1,2,…)の従う分布

$$f(x) = {}_{k+x-1}C_x\, p^k (1-p)^x \quad (x = 0,1,2,\cdots)$$



○○●○●○ …●
最後は成功なので，
$k+x-1$回からxの場所を決める組合せ数

二項分布で二項係数に負も認めた場合にこの分布になるので「負」の二項分布とよばれる

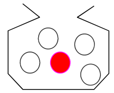


  - k=1のときは幾何分布に等しいため幾何分布の一般化となっている

その他の離散型分布

- 負の二項分布
  - 確率分布

$$f(x) = {}_{k+x-1}C_x\, p^k (1-p)^x,\ (x = 0,1,2,\cdots)$$



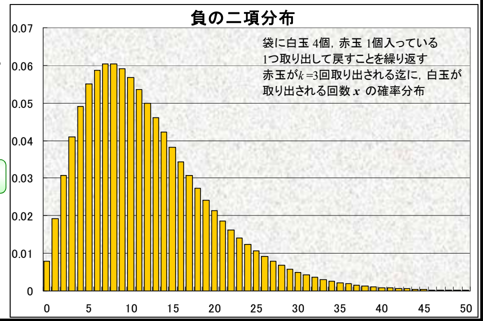
  - 期待値・分散

$$\begin{cases} E(X) = \dfrac{k(1-p)}{p}, \\ V(X) = \dfrac{k(1-p)}{p^2} \end{cases}$$

幾何分布のk倍

例では…

$$\begin{cases} E(X) = \dfrac{3(1-1/5)}{1/5} = 12.00, \\ V(X) = \dfrac{3(1-1/5)}{(1/5)^2} = 60.00 \end{cases}$$

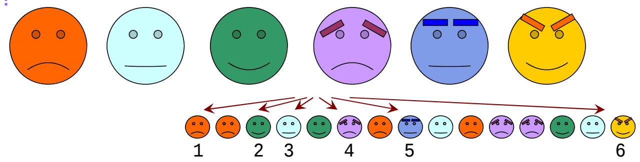


その他の離散型分布

- 負の二項分布の例
  - 例1：災害の到来
    - ある1年に風水害が起こる確率が1/25であるとする．風水害が起こるのは平均何年に1回か?
    - 上記と同じ災害が20年以内に起こる確率は?
  - 例2：袋から…
    - 白玉4つ，赤玉1つが入っている袋がある．1つ取り出して元に戻すという試行を繰り返したとき，赤玉が3回取り出されるまでに白玉が40回取り出される確率は?
  - 例3：シリーズものコレクター
    - 12種類のキャラクターが売られている．ただし，箱を開けるまで中にどれが入っているかはわからない．あるコレクターが全てのキャラクターを集めるためには何個買わねばならないか?

演習6

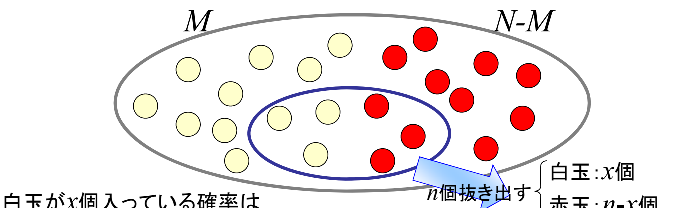
- 負の二項分布を求めよう
  - お菓子の付録に6種類のオマケがある．このオマケはそれぞれの箱に全種類がランダムに等確率に入っているとする．箱を開けるまで中身はわからない．黄色のオマケを3個そろえたい．そのために，お菓子を平均何個買わねばならないか?



注）全てのオマケを（少なくとも一つずつ）集める場合は，「クーポン収集問題」と呼ばれる問題になり，負の二項分布では計算できない．

その他の離散型分布

- 超幾何分布 hypergeometric distribution
  - 例：白玉がM個，赤玉がN-M個（全部でN個）ある．ここからn個抜き出したとき，白玉がx個入っている確率は?



白玉がx個入っている確率は…

$$f(x)=\frac{{}_M C_x \cdot {}_{N-M} C_{n-x}}{{}_N C_n}$$

n個取り出す組合せのうち，白玉x個，赤玉n-x個取り出す組合せの確率

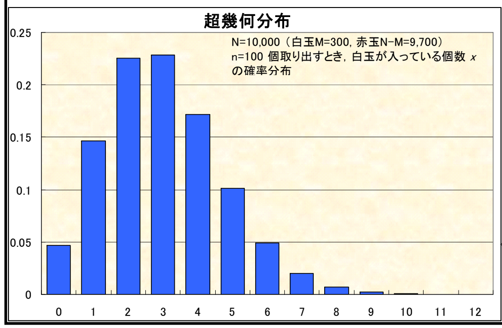
ただし，xの取り得る範囲は

$$(x=\max\{0,n-(N-M)\},\cdots,\min\{n,M\})$$

その他の離散型分布

- 超幾何分布
  - 確率分布

$$f(x)=\frac{{}_M C_x \cdot {}_{N-M} C_{n-x}}{{}_N C_n} \qquad (x=\max\{0,n-(N-M)\},\cdots,\min\{n,M\})$$



- 期待値・分散

$$\begin{cases} E(X)=\dfrac{Mn}{N}, \\ V(X)=\dfrac{Mn}{N}\cdot\dfrac{N-M}{N}\cdot\dfrac{N-M}{N-1} \end{cases}$$

例では…

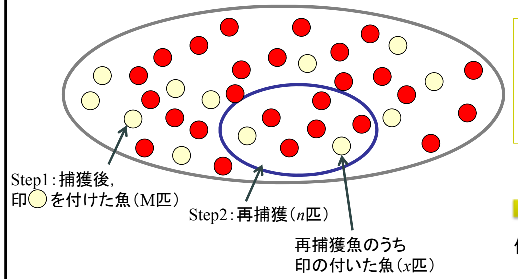
$$\begin{cases} E(X)=\dfrac{300\cdot 100}{10000}=3.00, \\ V(X)=\dfrac{300\cdot 100}{10000}\cdot\dfrac{9700}{10000}\cdot\dfrac{9700}{9999}=2.82 \end{cases}$$

Coffee Break!

標識再捕獲法
(mark-recapture method)
ともいう

「捕獲再捕獲法 capture-recapture method」

- 湖の中の魚の個体数（N匹）を推定したい
  - Step1（ランダムに）魚を捕獲（M匹），印 ○ をつけて放す
  - Step2：しばらくおいて，（ランダムに）魚を再捕獲（n匹）し，印の付いている魚を数える（x匹）



未知の推定したい数値

個体総数＝N匹
全○ M匹，全● N-M匹

再捕獲＝n匹
○ x匹，● n-x匹

既知の観測数値

推定値：$\hat{N} = \frac{Mn}{x}$

例）M=300, n=500, x=5なら

$$\hat{N} = \frac{300 \cdot 500}{5} = 30000$$

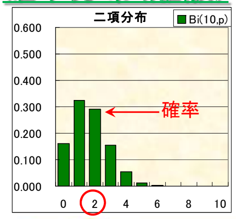
確率密度関数
p.d.f. (probability density function)

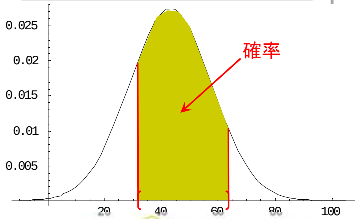
連続型分布continuous distribution

★正規分布 normal distribution
★標準正規分布 standard normal distribution
★χ^2分布 chi-square distribution
★t分布 Student's t distribution
★F分布 F distribution
★（連続型）一様分布 uniform distribution
★指数分布 exponential distribution

確率密度関数 *p.d.f.*

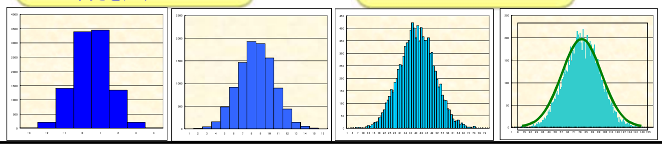
- 確率分布（離散） と 確率密度関数（連続）





各値（確率変数の値）に対し棒の高さが確率を表す
全ての高さを足すと1

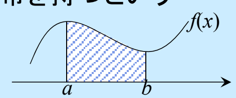
2変数に挟まれる部分の面積が確率を表す
全ての面積を足すと1



参考：確率密度関数 *p.d.f.*

- 連続（型）分布 continuous distribution
  - 確率変数 X の取る値が関数 $f(x)$ により，以下で与えられている場合，X は連続型の確率分布を持つという

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



ただし，

$$\begin{cases} f(x) \ge 0 (-\infty \le x \le \infty), \\ \int_{-\infty}^{\infty} f(x)dx = 1 \end{cases} \longleftrightarrow \quad f(x_k) := P(X = x_k) \ (k = 1,2,\cdots) \begin{cases} f(x_k) \ge 0 \ (k = 1,2,\cdots), \\ \sum_{k=1}^{\infty} f(x_k) = 1 \end{cases}$$

確率密度関数
probability density function

- 累積分布関数 c.d.f., cumulative distribution function

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt \longleftrightarrow F(x) = \sum_{u \le x} f(u)$$

参考：確率密度関数 *p.d.f.*

- 連続型確率変数の期待値と分散
  - 連続型確率変数 X の期待値

$$E(X)=\int_{-\infty}^{\infty} x\cdot f(x)dx \longleftrightarrow \sum_x x\cdot f(x)$$

  - 連続型確率変数 X の分散

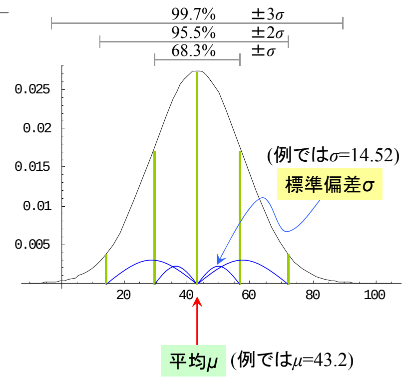
$$V(X)=\int_{-\infty}^{\infty} (x-E(X))^2 f(x)dx \longleftrightarrow \sum_x (x-E(X))^2 f(x)$$

1. 正規分布

- **正規分布 normal distribution** $N(\mu,\sigma^2)$
  - 確率密度関数

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

平均：$E(X)=\mu$
分散：$V(X)=\sigma^2$



1. 正規分布

- **標準正規分布 standard normal distribution** $N(0,1)$
  - 確率変数の標準化
    - 平均μ，分散σ^2の正規分布に従う確率変数Xを標準化する

$$X \to Z=\frac{X-\mu}{\sigma}$$

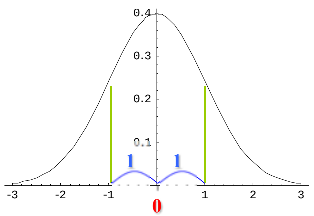
標準正規分布

⇨ 確率変数Zは，**平均0**，**分散1**の正規分布に従う

  - 確率密度関数

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

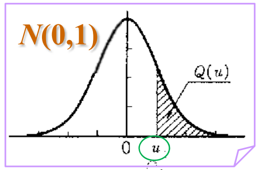
平均：$\mu=0$
分散：$\sigma^2=1$
（標準偏差：$\sigma=1$）



1. 正規分布

- **標準正規分布表とその読み方**
  - 標準正規分布表…分布の確率を表す一覧表

$$Q(u)=1-\Phi(u)=\int_u^{\infty}\phi(u)du$$



uの小数第2位

Xがu以上である確率 $P(X\geq u)$

uの小数第1位

u	.00	.01	.02	.03	.04	.05	.06
.0	.50000	.49601	.49202	.48803	.48405	.48006	.4760
.1	.46017	.45620	.45224	.44828	.44433	.44038	.4364
.2	.42074	.41683	.41294	.40905	.40517	.40129	.3974
.3	.38209	.37828	.37448	.37070	.36693	.36317	.3594
.4	.34458	.34090	.33724	.33360	.32997	.32636	.3227
.5	.30854	.30503	.30153	.29806	.29460	.29116	.2877
.6	.27425	.27093	.26763	.26435	.26109	.25785	.2546
.7	.24196	.23885	.23576	.23270	.22965	.22663	.2236
.8	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

1. 正規分布

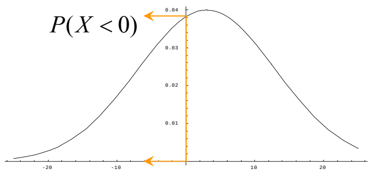
- **標準化と標準正規分布**
  - 例題：確率変数 X はある株式の利回り(%)で，正規分布 $N(3,10)$に従う．この株式への投資が損となる確率は？

$$P(X<0)$$
$$=P(\mu+\sigma Z<0)\ \left[\because Z=\frac{X-\mu}{\sigma}\right]$$
$$=P(Z<\frac{0-\mu}{\sigma})$$
$$=P(Z<\frac{0-3}{\sqrt{10}}=-0.94868)$$
$$\approx P(Z<-0.95)=0.17106$$

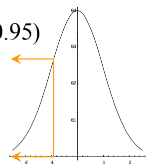
標準正規分布表から

＝0.171391 (Excel関数 NORMDISTより)

$P(X<0)$

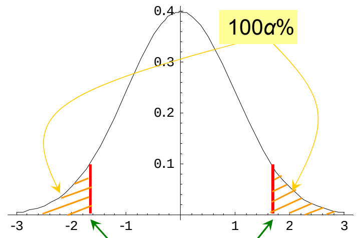


$P(Z<-0.95)$



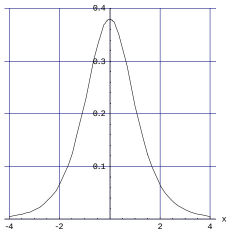
Coffee Break!

- パーセント点 percentage point
  - 分布の100α%点
    - 上側100α%点
    - 下側100α%点
    - 両側100α%点
  - 例：標準正規分布の両側100α%点
    - α=0.05 → 両側5%点
    - α=0.10 → 両側10%点
    - 標準正規分布$N(0,1)$の両側5%点（上側2.5%,下側2.5%）=±1.96



2. t 分布

- t分布 Student's t distribution



$t(v)$

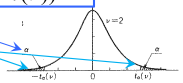
自由度v の t分布

自由度v は1,2,3,…の値をとり，自由度の値によって分布の形状が変わる

$$\alpha=P(T\geq t(v))=P(T\leq t(v))$$
$$2\alpha=P(T\geq t(v),T\leq t(v))$$

- t分布表の読み方

片側確率α
両側確率2α



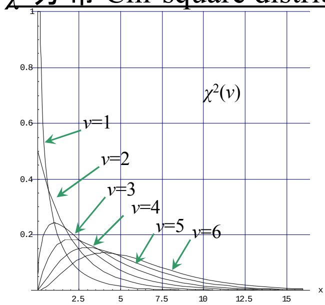
自由度v

$t(v)$

v \ α (2α)	.250 (.500)	.200 (.400)	.150 (.300)	.100 (.200)	.050 (.100)	.025 (.050)	.010 (.020)	.005 (.010)	.0005 (.0010)
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587

3. χ^2分布

- χ^2分布 Chi-square distribution



$\chi^2(v)$

自由度v の χ^2分布

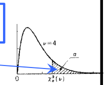
自由度v は1,2,3,…の値をとり，自由度の値によって分布の形状が変わる

$$\alpha=P(\chi^2\geq\chi^2(v))$$

- χ^2分布表の読み方

$\chi^2(v)$

確率α

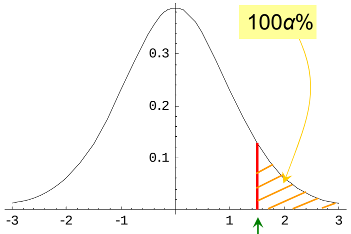


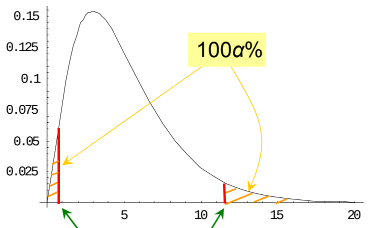
自由度v

v	.995	.990	.975	.950	.900	.800	.700	.600	.500
1	$.0^4392704$	$.0^3157088$	$.0^3982069$	$.0^2393214$	.0157908	.0641848	.148472	.274996	.45493
2	.0100251	.0201007	.0506356	.102587	.210721	.446287	.713350	1.02165	1.38629
3	.0717218	.114832	.215795	.351846	.584374	1.00517	1.42365	1.86917	2.36597
4	.206989	.297109	.484419	.710723	1.06362	1.64878	2.19470	2.75284	3.35669
5	.411742	.554298	.831212	1.14548	1.61031	2.34253	2.99991	3.65550	4.35146
6	.675727	.872090	1.23734	1.63538	2.20413	3.07009	3.82755	4.57015	5.34812
7	.989256	1.23904	1.68987	2.16735	2.83311	3.82232	4.67133	5.49323	6.34581
8	1.34441	1.64650	2.17973	2.73264	3.48954	4.59357	5.52742	6.42265	7.34412
9	1.73493	2.08790	2.70039	3.32511	4.16816	5.38005	6.39331	7.35703	8.34283
10	2.15586	2.55821	3.24697	3.94030	4.86518	6.17908	7.26722	8.29547	9.34182
11	2.60322	3.05348	3.81575	4.57481	5.57778	6.98867	8.14787	9.23729	10.3410
12	3.07382	3.57057	4.40379	5.22603	6.30380	7.80733	9.03428	10.1820	11.3403
13	3.56503	4.10692	5.00875	5.89186	7.04150	8.63386	9.92568	11.1291	12.3398
14	4.07467	4.66043	5.62873	6.57063	7.78953	9.46733	10.8215	12.0785	13.3393
15	4.60092	5.22935	6.26214	7.26094	8.54676	10.3070	11.7212	13.0297	14.3389

Coffee Break!

- パーセント点 percentage point
  - 例：自由度10のt分布の上側100α%点
  - 例：自由度5のχ^2分布の両側100α%点





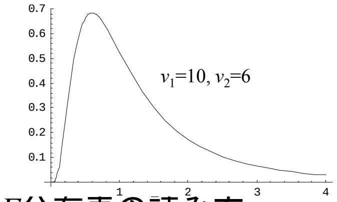
4. F分布

- F分布 F distribution

$\boxed{F(\nu_1, \nu_2)}$

自由度ν_1, ν_2の F分布

自由度ν_1, ν_2 は各々1,2,3,…の値をとり，自由度の値によって分布の形状が変わる



- F分布表の読み方

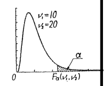
$$\boxed{\alpha = P(F \geq F(\nu_1, \nu_2))}$$

確率α → $\alpha=0.05$

自由度ν_1

$F(\nu_1, \nu_2)$

自由度ν_2



ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020

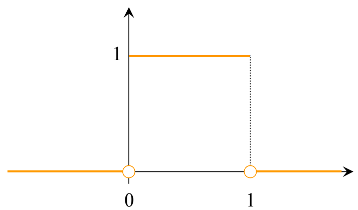
その他の連続型分布

- （連続型）一様分布 uniform distribution
  - 確率密度関数

$$f(x) = \begin{cases} 1 & [0,1] \\ 0 & o.w. \end{cases}$$

  - 累積分布関数

$$F(x) = \begin{cases} 0 & (x < 0) \\ x & (0 \leq x \leq 1) \\ 0 & (1 < x) \end{cases}$$



  - 期待値・分散

$$\begin{cases} E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_0^1 x \cdot 1dx = \left[\frac{x^2}{2}\right]_0^1 = \left\{\frac{1^2}{2} - \frac{0^2}{2}\right\} = \frac{1}{2} \\ V(X) = \int_{-\infty}^{\infty} (x - E(X))^2 \cdot f(x)dx = \int_0^1 (x - \frac{1}{2})^2 \cdot 1dx = \left[\frac{(x-1/2)^3}{3}\right]_0^1 \\ = \left\{\frac{(1-1/2)^3}{3} - \frac{(0-1/2)^3}{3}\right\} = \frac{1/8}{3} - \frac{-1/8}{3} = \frac{1}{12} \end{cases}$$

その他の連続型分布

- 指数分布 exponential distribution $\boxed{Ex(\lambda)}$

ポアソン分布に従って起きる事象の生起間隔を表現

  - 確率密度関数

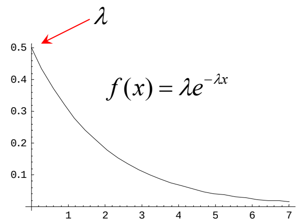
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & o.w. \end{cases}$$

  - 累積分布関数

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & (x \geq 0) \\ 0 & o.w. \end{cases}$$

  - 期待値・分散

$$E(X) = \frac{1}{\lambda}, \ V(X) = \frac{1}{\lambda^2}$$

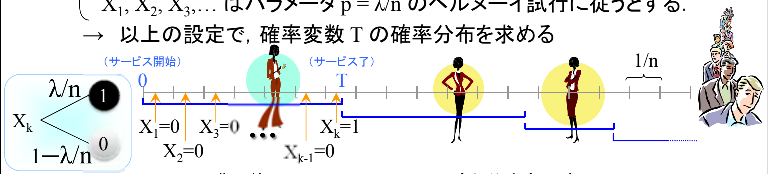


その他の連続型分布

- 指数分布 exponential distribution
  - 例：サービスの待ち時間（チケット売場の列）

単位時間をn等分し，ある区間kで誰かが購入を終了したら $X_k=1$，そうでないなら $X_k=0$ とする確率変数 $X_1, X_2, X_3,\dots$ を考える．
チケット販売開始時点 = 0, 1人目の購入終了時点 = T とする．
$X_1, X_2, X_3,\dots$ はパラメータ $p=\lambda/n$ のベルヌーイ試行に従うとする．

→ 以上の設定で，確率変数 T の確率分布を求める



区間N+1で購入終了 → T≒(1/n)N （nが十分大きい時）
任意の t>0に対し，kを，k≦nt<k+1を満たす整数とする
N+1は幾何分布に従う

$$\to P(0 \le T \le t) \approx P(0 \le N \le nt) = \sum_{i=0}^{k}(1-\lambda/n)^i \lambda/n$$

$$= 1-(1-\lambda/n)^{k+1} \approx 1-(1-\lambda/n)^{nt} \xrightarrow[n\to\infty]{} 1-e^{-\lambda t} = \int_0^t \lambda e^{-\lambda x}dx$$

累積分布関数

確率密度関数

参考：その他の連続型分布

- ガンマ分布 Gamma distribution　$Ga(\alpha,\lambda)$
  - 確率密度関数

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x} \quad (x \ge 0)$$

  - ガンマ関数 Gamma function

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx \quad (\alpha > 0)$$

$$\Gamma(n) = (n-1)! \qquad (n \in N)$$

nが自然数の時

$\Rightarrow$ $\begin{cases} Ga(n/2, 1/2) & \text{：自由度 n の}\chi^2\text{分布} \\ Ga(1,\lambda) & \text{：指数分布} \end{cases}$

参考：その他の連続型分布

- ベータ分布 Beta distribution　$Be(\alpha, \beta)$
  - 確率密度関数

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \quad (0 < x < 1)$$

  - ベータ関数 Beta function

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \quad (\alpha,\beta > 0)$$

$$= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

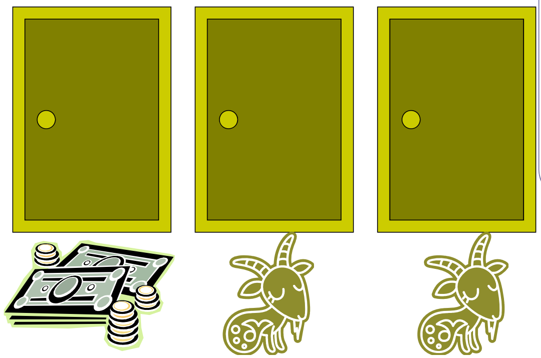
Coffee Break!

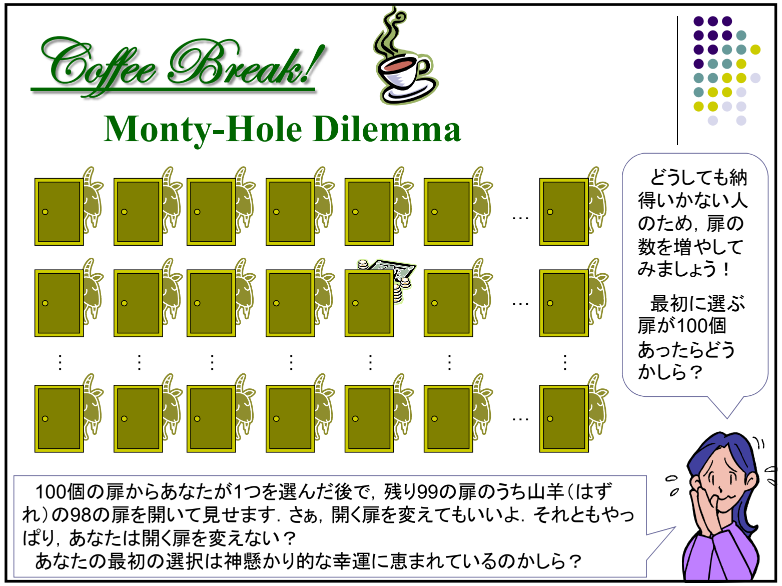
Monty-Hole Dilemma
確率的直感

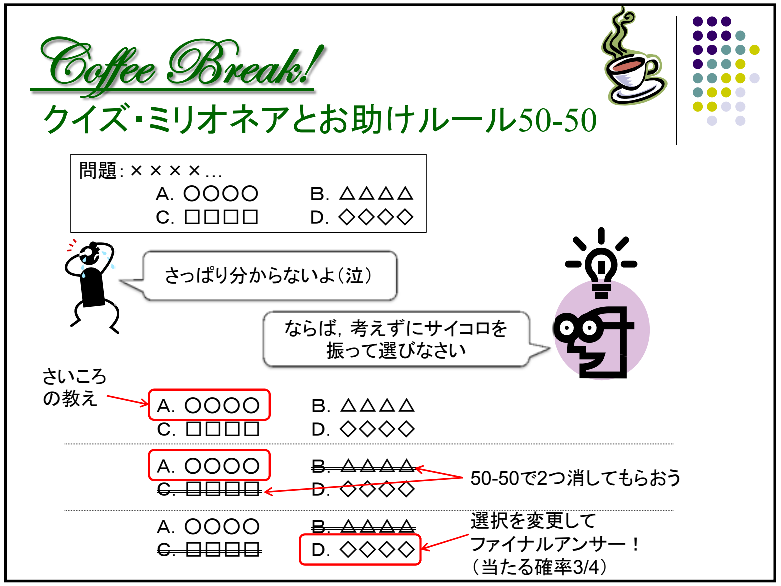
3枚の扉の向こうに
•百万ドル（当たり）
•山羊（はずれ）
•山羊（はずれ）
が隠されているよ．
あなたは扉を1つだけ選んでいいのよ．

ところで，あなたが選ばなかった2つの扉のうち，山羊の扉を開くから，それを見た後で，開く扉を変えてもいいよ．

さぁ，どうする？







参考文献

- 東京大学教養学部統計学教室編 「統計学入門」 東京大学出版会(1991)
- 村上雅人「なるほど統計学」 海鳴社 (2002)
- 金子治平ほか「よくわかる統計学 I 基礎編」ミネルヴァ書房(2007)
- 田栗正章ほか「やさしい統計入門」講談社(2007)
- G.Blom, L.Holst, D.Sqndell / 森真 訳「確率論へようこそ」シュプリンガー・フェアラーク東京(1995,2005[新装版])

- 丹慶勝市「図解雑学 統計解析」ナツメ社 (2003)
- 東京大学教養学部統計学教室編 「自然科学の統計学」 東京大学出版会(1992)
- 白石修二「例題で学ぶ Excel統計入門」 森北出版(2001)
- J.Matousek, J.Nesetril / 根上生也・中本敦浩 訳「離散数学への招待 上」シュプリンガー・フェアラーク東京(2002)
- 徳山豪「工学基礎 離散数学とその応用」数理工学社(2003)
- B.Schechter / グラベルロード訳「My Brain is Open」共立出版(2003)