物理実験学 I

2012 年度　前期

講義メモ ver20120525

（身内担当分・オリジナルは蔵重）

http://ppwww.phys.sci.kobe-u.ac.jp/~miuchi

から　講義ノート、実験テキストなど

日程

		内容	対応章
第 1 講	4/13	確率・統計の概念、確率変数と確率分布	0,1-1～1-3
第 2 講	4/20	期待値、分散、ベルヌーイ試行	1-4～1-5
第 3 講	4/27	ポアソン分布、正規分布	1-6～2-1
第 4 講	5/11	確率の法則、標本、母集団	2-2～3-2
第 5 講	5/18	推定、最尤法、最小二乗法	3-3～3-7
第 6 講	5/25	検定	4-1～4-3
第 7 講	6/1	中間試験	

後半日程（担当：松岡）：　6/8、6/15、6/22、6/29、7/6、7/13(休講)、7/20、7/27（期末試験）

連絡先　自然科学総合研究棟 3 号館 317 号室　/　miuchi@panda.kobe-u.ac.jp

参考書　岩波書店　理工系の基礎数学 7　確率・統計　柴田文明

この講義の目指すところ

●物理実験で得られた結果を正しく解釈する。

⇔ある仮説は（どれくらいの確率で）棄却されるか?

● 世の中の「統計」を正しく解釈する。

●実験はおもしろい。実験学は　どうだ？

改定 LOG

20120412：第二講まで整理。

20120419：第二講まで再度整理。

20120426：第四講まで整理。

20120511：第四講まで再度整理。

20120525：第六講まで整理。例題を追加。

20120803:まとめ地図のポアソン分布の分散=μ^2のミスを分散=μに修正

<<第１講>>

０： Introduction

記述統計（*descriptive statistics*）

記述統計とは、収集したデータの要約統計量（平均、分散など）を計算して分布を明らかにする事により、データの示す傾向や性質を知ること。

例： 試験の成績、 国勢調査、

数理統計学・推計統計（*mathematical statistics*）

データからその元となっている諸性質を確率論的に推測する分野。

例： 標本調査 ： 全数調査が不可能であるとき

降水率 ： 予測

推定：標本から母集団の統計学的な性質を推測すること

検定：母集団が特定の分布に従っているかどうかを検証すること

例；乾電池の寿命

標本を選らんで、寿命を調査→平均寿命、分散を求める

→分布を仮定すると、任意の１個がある寿命以下の確率が求まる

例：放射線物質の含有量

一定時間に計測された放射線の数を測定

→含有量が統計誤差で求まる（推定）

*

間違った推定・検定

- 横断歩道での交通事故記録の統計をとったところ、青信号の横断中の方が、赤信号の横断中よりも、交通死亡事故が多いという統計結果が得られた。青信号で渡った方が事故にあう危険が高い。

  ←母集団における分布をとりいれていない場合
- 小数被験者の実験で結論を出してしまうテレビ番組の検証実験

  ←統計誤差が考慮されていない場合
- インターネットによるアンケート調査の統計

  ←標本が無作為抽出されていない場合
- 成功した例しか掲載しないダイエット広告

  ←作為的なデータの棄却、選択。

１：　基礎的な事柄

1-1 事象・集合・確率

〇基礎概念の定義

・試行：　ある操作

・標本点(sample point)：　起こりえる可能な結果

・標本空間(sample space)：標本点の全体の集合　Ω

・事象(event)：　起こりえる事柄。標本空間の部分集合

・空事象(empty event)：　標本点を一つも含まない起こりえない事柄　ϕ

・根元事象(elementary event)：　ただ一つの標本点からなる事象

・複合事象(composite event)：　複数の標本点を含む事象

例：さいころ

試行：　さいころを振って出目をみる

標本点:　‘1’,’2’,’3’,’4’,’5’,’6’　（‘1’さいころの目１がでること）

標本空間:｛‘1’,’2’,’3’,’4’,’5’,’6’｝

空事象:‘7’

根元事象：　｛‘1’},{’2’　},{’3’　},{’4’　},{’5’　},{’6’｝

複合事象：　{奇数}＝｛‘1’,’3’,’5’｝

・和事象(union of events)：$A \cup B$

　　AとＢの少なくともどちらか一つが起こるという事象

・積事象(intersection of events)：$A \cap B$

　　ＡとＢ両方が同時に起こるという事象

・補事象(complementary event)：$\overline{\mathrm{A}}$

　　事象Aが起こらないという事象　$\overline{\mathrm{A}} \cap \mathrm{A} = \phi$　$\bar{A} \cup \mathrm{A} = \Omega$

・排反事象(disjoint events)：事象ＡとＣが同時に起こりえないとき，ＡとＣは排反事象

$A \cap C = \phi$

・分配法則

$$(\mathrm{A} \cup \mathrm{B}) \cap \mathrm{C} = (\mathrm{A} \cap \mathrm{C}) \cup (\mathrm{B} \cap \mathrm{C}) \tag{1.1.1}$$

$$(\mathrm{A} \cap \mathrm{B}) \cup \mathrm{C} = (\mathrm{A} \cup \mathrm{C}) \cap (\mathrm{B} \cup \mathrm{C}) \tag{1.1.2}$$

・ドモルガンの法則

$$\overline{(\mathrm{A} \cap \mathrm{B})} = \overline{\mathrm{A}} \cup \overline{\mathrm{B}} \tag{1.1.3}$$

$$\overline{(\mathrm{A} \cup \mathrm{B})} = \overline{\mathrm{A}} \cap \overline{\mathrm{B}} \tag{1.1.4}$$

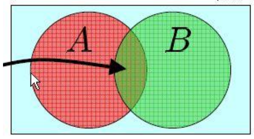
○確率

・確率（probability）：事象の起こりやすさを定量的に示すもの

どの根源事象も同じ確からしさでおこるとき、事象Ａの起こる確率　$P(A) = \frac{n(A)}{n(\Omega)}$

・コルモゴロフ(Kolmogorov)　の公理

1. 任意の事象に対して$0 \leq P(A) \leq 1$　　$P(\phi) = 0$　$P(\Omega) = 1$　(1.1.5)
2. 全確率は１：$P(\Omega) = 1$　(1.1.6)
3. 互いに排反な事象A_iに対して$P(\cup A_i) = \sum P(A_i)$　(1.1.7)



・加法定理

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$　(1.1.8)

・条件付き確率

事象Ｂが起こったとわかっている場合に事象Ａが起こる確率を，Ｂを条件とするＡの条件付き確率（conditional probability）とよぶ

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.1.9)$$

・乗法定理

$$\boxed{P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)} \quad (1.1.10)$$

・独立性

事象ＡとＢが$P(A \cap B) = P(A)P(B)$を満たすとき、ＡとＢは独立（independent）

つまり$P(A|B) = P(A), P(B|A) = P(B)$

例：

A：｛さいころを振って奇数が出るという事象｝　＝｛‘1’，’3’，’5’｝　$P(A) = \frac{3}{6}$

B：｛さいころを振って３以下の目が出るという事象｝＝｛‘1’，’2’，’3’｝　$P(B) = \frac{3}{6}$

A∩B　＝｛‘1’，’3’｝　$P(A \cap B) = \frac{2}{6}$

A∪B　＝｛‘1’，’2’　，’3’，’5’｝　$P(A \cup B) = \frac{4}{6}$　　$P(A) + P(B) - P(A \cap B) = \frac{3}{6} + \frac{3}{6} - \frac{2}{6} = \frac{4}{6}$

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{3}$

$P(A|B)P(B) = \frac{2}{3}\frac{3}{6} = \frac{2}{6}$

**

1-2 順列・組合せ

・順列（permutation）：ｎ個の異なるものから、ｒ個のものを取り出して並べたもの

順列の総数は

$$nPr = n \cdot (n-1) \cdot (n-2) \cdots\cdots (n-r+1) = \frac{n!}{(n-r)!} \qquad (1.2.1)$$

$$n! = 1 \times 2 \times \ldots \times n \quad (\text{階乗}),\quad 0! = 1 \qquad (1.2.2)$$

・組み合わせ（combination）：ｎ個の異なるものから、ｒ個のものを取り出したもの(順番は気にしない)

組合せの総数は

$$nCr = \frac{nPr}{r!} = \frac{n!}{r! \cdot (n-r)!} \qquad (1.2.3)$$

例：　超幾何分布

袋の中にＡが M 個，Ｂが $N-M$ 個，合計 N 個の玉が入っている．

この袋から無作為に玉を n 個取り出したときのＡの数を x 個，Ｂの数を $n-x$ 個である確率は

$$f(x) = \frac{{}_M C_x \times {}_{N-M} C_{n-x}}{{}_N C_n} \qquad (1.2.4)$$

1-3 確率変数・確率分布

〇定義

・確率変数（random variable）：とる値に対して確率が与えられている変数
・実現値：確率変数が実際にとる値
　確率変数は大文字で，実現値は小文字で表わすことが多い

・確率分布（probability distribution）：確率変数の実現値と確率の関係を関数として表現したもの

$$P(X=x)=f(x) \tag{1.3.1}$$

・累積分布関数(cumulative distribution function)：確率変数がある値以下をとる確率

$$F(x)=P(X\leq x) \tag{1.3.2}$$

〇離散型(discrete type)確率変数：可算集合の中の値をとる確率変数

例：

確率変数X:　さいころを振ったときの出目の値

実現値x : 1, 2, 3, 4, 5, 6

$$f(x)=\frac{1}{6}, x=1,2,3,4,5,6 \qquad F(x)=\frac{x}{6}$$

〇連続型(continuous type)確率変数：連続値をとる確率変数

・連続型の確率変数の確率分布：:確率密度関数(probability density function)

$$P(\mathrm{x}\leq X\leq x+dx)=f(x)dx \tag{1.3.3}$$

$$P(a\leq X\leq b)=\int_{\mathrm{a}}^{\mathrm{b}}f(x)dx \tag{1.3.4}$$

$$P(-\infty\leq X\leq\infty)=\int_{-\infty}^{\infty}f(x)dx=1 \tag{1.3.5}$$

・累積分布関数(cumulative distribution function)：確率変数がある値以下をとる確率

$$F(x)=P(X\leq x)=\int_{-\infty}^{x}f(y)dy \tag{1.3.6}$$

$$P(a\leq X\leq b)=\int_{\mathrm{a}}^{\mathrm{b}}f(x)dx=F(b)-F(a) \tag{1.3.7}$$

例：

確率変数X:　時計の分針の角度

実現値x : $0\leqq x<360$

$$f(x)=\frac{1}{360}$$

$$F(x)=P(X\leq x)=\int_{0}^{x}\frac{1}{360}dy=\frac{x}{360}$$

$$P(a\leq X\leq b)=F(b)-F(a)=\frac{b-a}{360}$$

<<第 1 講ここまで>>

<<第 2 講>>

1-4 期待値と分散： 確率変数の性質を表わす指標

○期待値

・期待値(expectation)： 確率変数の値の平均（正確には確率による重み付きの平均）

確率変数Xの期待値を$E(X)$ または ⟨X⟩ で表す

離散型：$E(X) = \sum x \cdot f(x)$ (1. 4. 1)

連続型：$E(X) = \int x \cdot f(x)dx$ (1. 4. 2)

以後，説明を簡単にするため，連続型の確率変数を主に扱うことにする．離散型を考える場合は，積分を和に変更すればよい．

・期待値演算の性質

定数は期待値をとっても値は変わらない

$$E(c) = \int c \cdot f(x)dx = c\int f(x)dx = c \quad (1.4.3)$$

定数を足した期待値は，期待値に定数を足したものと等しい

$$E(X + c) = \int (x + c) \cdot f(x)dx = E(X) + c \quad (1.4.4)$$

定数倍の期待値は，期待値の定数倍と等しい

$$E(cX) = \int cx \cdot f(x)dx = c\int x \cdot f(x)dx = cE(X) \quad (1.4.5)$$

→期待値演算は線形

○分散

分散(variance)：確率変数の散らばり具合を表す指標。

確率変数Xの分散V(X)

$$\boxed{V(X) = E[\{X - E(X)\}^2]} \quad (1.4.6)$$

$$V(X) = E[\{X - E(X)\}^2] = E\left[X^2 - 2\mathrm{X} \cdot \mathrm{E(X)} + \left(\mathrm{E(X)}\right)^2\right]$$

$$= \int \left[x^2 - 2x \cdot E(X) + \left(\mathrm{E(X)}\right)^2\right] f(x)dx$$

$$= \int x^2 f(x)dx - 2E(X)\int x\,f(x)dx + \left(\mathrm{E(X)}\right)^2 \int f(x)dx$$

$$= E[X^2] - 2\mathrm{E(X)} \cdot E[\mathrm{X}] + \left(\mathrm{E(X)}\right)^2 \cdot E[1]$$

$$\boxed{= E[X^2] - \left(\mathrm{E(X)}\right)^2} \quad (1.4.7)$$

・分散演算の性質

定数の分散はゼロ

$$V(c) = E[\{c - E(c)\}^2] = E[0] = 0 \quad (1.4.8)$$

定数を足したものの分散は，もとの分散と等しい

$$V(X+c)=E[\{X+c-E(X+c)\}^2]=E[\{X-E(X)\}^2]=V(X) \tag{1.4.9}$$

定数倍の分散は，もとの分散に定数の２乗をかけたものと等しい

$$V(cX)=E[\{cX-E(cX)\}^2]=E[c^2\{X-E(X)\}^2]=c^2V(X) \tag{1.4.10}$$

〇標準偏差

標準偏差(standard deviation)：分散の平方根

$$\boxed{D(X)=\sqrt{V(X)}} \tag{1.4.11}$$

分散の値をσ２で，標準偏差の値をσで表わすことが多い

〇標準化(standardization)

任意の確率変数 X に対して

$$Z=\frac{X-E(X)}{D(X)} \tag{1.4.12}$$

と定義すれば，Z は期待値０，分散１になる

$$E(Z)=\frac{E(X-E(X))}{D(X)}=\frac{E(X)-E(X)}{D(X)}=0 \tag{1.4.13}$$

$$V(Z)=\frac{V(X-E(X))}{\{D(X)\}^2}=\frac{V(X)}{V(X)}=1 \tag{1.4.14}$$

期待値 50、分散 10 に規格化したものを一般に「偏差値」と呼んでいる。

例 1-4-1：

さいころを１つふったときの期待値と分散

$$\sum_{N=1}^{6}N=21,\quad \sum_{N=1}^{6}N^2=91,\quad \sum_{N=1}^{6}N^3=441$$

$$\langle X\rangle=\sum_{i=1}^{6}i\times\frac{1}{6}=\frac{7}{2}$$

$$\sigma^2{}_X=\langle X^2\rangle-\langle X\rangle^2=\frac{1}{6}\sum_{i=0}^{6}i^2-\left(\frac{1}{6}\sum_{i=0}^{6}i\right)^2=\frac{91}{6}-\left(\frac{7}{2}\right)^2=\frac{182-147}{12}=\frac{35}{12}$$

○確率変数の和

確率変数X とY が独立である、つまり $f(X,Y)=g(X)h(Y)$であるとき、

その和 $Z=X+Y$の期待値$E(Z)$は

$$\boxed{E(Z)}=E(X+Y)=\iint(x+y)f(X,Y)dxdy$$

$$=\iint(x+y)g(x)h(y)dxdy$$

$$=\int xg(x)dx\int h(y)dy+\int g(x)dx\int yh(y)dy=\boxed{E(X)+E(Y)} \quad (1.4.15)$$

Zの分散$V(Z)$は(1.4.7)より

$$\boxed{V(Z)}=E[(X+Y)^2]-\{\mathrm{E(X+Y)}\}^2 \quad (1.4.16)$$

$$E[(X+Y)^2]=\iint(x+y)^2f(X,Y)dxdy=\iint(x+y)^2g(x)h(y)dxdy$$

$$=\int x^2g(x)dx\int h(y)dy+2\int xg(x)dx\int yh(y)dy+\int g(x)dx\int y^2h(y)dy$$

$$=E(X^2)+2E(X)E(Y)+E(Y^2)$$

$$V(Z)=\{E(X^2)+2E(X)E(Y)+E(Y^2)\}-\{\mathrm{E(X)+E(Y)}\}^2$$

$$=E(X^2)-(E(X))^2+E(Y^2)-(E(Y))^2=\boxed{V(X)+V(Y)} \quad (1.4.17)$$

○確率変数の積

確率変数X とY が独立である、つまり $f(X,Y)=g(X)h(Y)$であるとき、

その積 $Z=XY$の期待値$E(Z)$は

$$\boxed{E(Z)}=E(XY)=\iint(xy)f(X,Y)dxdy=\iint(xy)g(x)h(y)dxdy$$

$$=\int xg(x)dx\int yh(y)dy=E(X)E(Y) \quad (1.4.18)$$

となる。

Zの分散$V(Z)$は

$$\boxed{V(Z)}=E[(XY)^2]-\{\mathrm{E}(XY)\}^2 \quad (1.4.19)$$

$$E[(XY)^2]=\iint(xy)^2f(X,Y)dxdy=\int x^2g(x)dx\int y^2h(y)dy=E(X^2)E(Y^2)$$

$$=\{V(X)+(E(X))^2\}\{V(Y)+(E(Y))^2\} \quad (1.4.20)$$

$$V(Z)=\{E(X^2)E(Y^2)\}-\{E(X)E(Y)\}^2$$

$$=V(X)V(Y)+V(X)\big(E(Y)\big)^2+V(Y)\big(E(X)\big)^2$$

$$=\boxed{(E(Z))^2\left\{\frac{V(X)}{(E(X))^2}+\frac{V(Y)}{(E(Y))^2}+\frac{V(X)}{(E(X))^2}\frac{V(Y)}{(E(Y))^2}\right\}} \quad (1.4.21)$$

相対誤差$V(X)/\big(E(X)\big)^2$が小さいとき第三項を無視できて

$$\frac{V(Z)}{(E(Z))^2}\cong\left\{\frac{V(X)}{(E(X))^2}+\frac{V(Y)}{(E(Y))^2})\right\} \quad (1.4.21)$$

○誤差の伝播

確率変数X とY が独立である、つまり $f(X,Y)=g(X)h(Y)$であるときの、X, Yの関数 $Z(X,Y)$の期待値と分散を考える。

相対誤差$V(X)/\left(E(X)\right)^2$、$V(Y)/\left(E(Y)\right)^2$が小さいとき(つまり x,y とも期待値の近くに分布する)

$Z=Z(X,Y)$を$Z(\bar{x}=E(X),\bar{y}=E(Y))$のまわりで Taylor 展開して

$$Z(x,y)\cong Z(\bar{x},\bar{y})+\frac{\partial Z}{\partial x}(x-\bar{x})+\frac{\partial Z}{\partial y}(y-\bar{y}) \tag{1.4.22}$$

$E(Z)$は

$$E(Z)\cong\iint\left\{Z(\bar{x},\bar{y})+\frac{\partial Z}{\partial x}(x-\bar{x})+\frac{\partial Z}{\partial y}(y-\bar{y})\right\}g(x)h(y)dxdy=Z(\bar{x},\bar{y}) \tag{1.4.23}$$

分散 $V(Z)$は

$$V(Z)=\iint\{Z(x,y)-E(Z)\}^2f(X,Y)dxdy=\iint\left\{\frac{\partial Z}{\partial x}(x-\bar{x})+\frac{\partial Z}{\partial y}(y-\bar{y})\right\}^2g(x)h(y)dxdy$$

$$=\left(\frac{\partial Z}{\partial x}\right)^2\int(x-\bar{x})^2g(x)dx+\left(\frac{\partial Z}{\partial y}\right)^2\int(y-\bar{y})^2h(y)dy+\left(\frac{\partial Z}{\partial x}\right)\left(\frac{\partial Z}{\partial y}\right)\int(x-\bar{x})g(x)dx\int(y-\bar{y})\,h(y)dy$$

$$=\left(\frac{\partial Z}{\partial x}\right)^2V(X)+\left(\frac{\partial Z}{\partial y}\right)^2V(Y) \tag{1.4.24}$$

例 1-4-2：
さいころを２つふったときの平均値の、期待値と分散

$$E\left(\frac{X+Y}{2}\right)=\frac{E(X)+E(Y)}{2}=E(X)=3.5$$

$$V\left(\frac{X+Y}{2}\right)=\frac{V(X)+V(Y)}{4}=\frac{V(X)}{2}=\frac{35}{24}$$

例 1-4-3：
さいころを２つふったとき、大きい目から小さい目を引いたときの値 Zの、期待値と分散
同じ目が出るのは６通り
差が z (1≦z≦5) になるのは、2(6−z)通り

$$\sum_{z=0}^{5} z=15 \sum_{z=0}^{5} z^2=55 \sum_{z=0}^{5} z^3=225$$

$$E(Z)=\sum_{z=0}^{5} P(z)z=\sum_{z=1}^{5}\frac{2(6-z)}{36}z=\frac{1}{3}\sum_{z=1}^{5} z-\frac{1}{18}\sum_{z=1}^{5} z^2=\frac{15}{3}-\frac{55}{18}=\frac{35}{18}$$

$$V(Z)=E(Z^2)-\left(E(Z)\right)^2=\sum_{z=0}^{5} P(z)z^2-\left(\frac{35}{18}\right)^2=\sum_{z=1}^{5}\frac{2(6-z)}{36}z^2-\left(\frac{35}{18}\right)^2$$

$$=\frac{1}{3}\sum_{z=1}^{5} z^2-\frac{1}{18}\sum_{z=1}^{5} z^3-\left(\frac{35}{18}\right)^2$$

$$=\frac{55}{3}-\frac{225}{18}-\left(\frac{35}{18}\right)^2=\frac{5}{18^2}[1188-810-245]=\frac{665}{324}$$

*

1-5 ベルヌーイ試行と二項分布

○二項分布

ベルヌーイ試行(Bernoulli trials)：２つの根源事象しかない標本空間を考える。一回の試行で一方の根源事象 A が起こる確率を p とする。この実験を同じ条件で独立に回繰り返すこと。

n 回のベルヌーイ試行に対して，x 回事象 A が起こる確率は

$$\boxed{f(x) = {}_nC_x \cdot p^x \cdot (1-p)^{n-x}} \tag{1.5.1}$$

これを二項分布(Binomial distribution)とよび，$B(n, p)$で表す

・二項分布の全確率

$\sum_{x=0}^{n} f(x) = 1$を示す。

二項定理より

$$(p+q)^{\mathrm{n}} = \sum_{\mathrm{x=0}}^{\mathrm{n}} {}_nC_x \cdot p^x \cdot q^{n-x} \tag{1.5.2}$$

$q = 1 - p$を代入して

$$1 = (p+1-p)^{\mathrm{n}} = \sum_{\mathrm{x=0}}^{\mathrm{n}} {}_nC_x \cdot p^x \cdot (1-p)^{n-x} \tag{1.5.3}$$

$$\therefore \sum_{x=0}^{n} f(x) = \sum_{x=0}^{n} {}_nC_x \cdot p^x \cdot (1-p)^{n-x}(x) = 1 \tag{1.5.4}$$

・二項分布の期待値

1 回ごとの確率×回数　直感的に理解しやすい。

$$E(X) = \sum_{x=0}^{n} x f(x) = \sum_{x=0}^{n} x\, {}_nC_x \cdot p^x \cdot (1-p)^{n-x} = \boxed{\mathrm{np}} \tag{1.5.5}$$

これを証明するために、準備として二項定理の両辺を p で微分する。

$$\frac{\partial}{\partial \mathrm{p}}(p+q)^{\mathrm{n}} = \frac{\partial}{\partial \mathrm{p}} \sum_{\mathrm{x=0}}^{\mathrm{n}} {}_nC_x \cdot p^x \cdot q^{n-x} \tag{1.5.6}$$

$$n(p+q)^{\mathrm{n}-1} = \sum_{\mathrm{x=0}}^{\mathrm{n}} {}_nC_x \cdot x \cdot p^{x-1} \cdot q^{n-x} \tag{1.5.7}$$

この右辺に p を掛け、$q = 1 - p$を代入すると求めるものとなるので、

$$E(X) = \sum_{x=0}^{n} x\, {}_nC_x \cdot p^x \cdot (1-p)^{n-x} = p \cdot n(p+1-p)^{n-1} = np \tag{1.5.8}$$

・二項分布の分散

$$V(X) = E(X^2) - \{E(X)\}^2 = \boxed{np(1-p)} \tag{1.5.9}$$

$$E(X^2) = \sum_{x=0}^{n} x^2 f(x) = \sum_{x=0}^{n} x^2\, {}_nC_x \cdot p^x \cdot (1-p)^{n-x} \tag{1.5.10}$$

これを証明するために、準備として二項定理の両辺を p で 2 階微分する。

$$\frac{\partial^2}{\partial \mathrm{p}^2}(p+q)^{\mathrm{n}} = \frac{\partial^2}{\partial \mathrm{p}^2} \sum_{\mathrm{x=0}}^{\mathrm{n}} {}_nC_x \cdot p^x \cdot q^{n-x} \tag{1.5.11}$$

$$n(n-1)(p+q)^{\mathrm{n}-2} = \sum_{\mathrm{x=0}}^{\mathrm{n}} {}_nC_x \cdot x(x-1) \cdot p^{x-2} \cdot q^{n-x} \tag{1.5.12}$$

よって

$$E(X^2) = \sum_{x=0}^{n} x^2 f(x) = \sum_{x=0}^{n} x^2\, {}_nC_x \cdot p^x \cdot (1-p)^{n-x}$$

$$= p^2 \sum_{x=0}^{n} (x^2 - x)\, {}_nC_x \cdot p^{x-2} \cdot (1-p)^{n-x} + p \sum_{x=0}^{n} x\, {}_nC_x \cdot p^{x-1} \cdot (1-p)^{n-x}$$

$$= p^2\{n(n-1)(p+1-p)^{n-2}\} + p\{n(p+1-p)^{n-1}\} = n(n-1)p^2 + np \tag{1.5.13}$$

$$V(X) = E(X^2) - \{E(X)\}^2 = n(n-1)p^2 + np - (np)^2 = np(1-p) \tag{1.5.14}$$

**

<<第 2 講ここまで>>

<<第３講>>

1-6 ポアッソン分布

・ポアソン分布の定義

ベルヌーイ試行に対して，事象 A が起こる回数の期待値$\mu \equiv np$を固定して

$n \to \infty, p \to 0$とした極限をとってみよう

二項分布$B(n,p)$をμを用いて書き換えて

$$f(x) = \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} \cdot \left(\frac{\mu}{n}\right)^x \cdot \left(1-\frac{\mu}{n}\right)^{n-x}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-x+1}{n} \frac{\mu^x}{x!} \left(1-\frac{\mu}{n}\right)^{n-x} \tag{1.6.1}$$

$n \gg x$であるから、

$$f(x) \cong \frac{\mu^x}{x!} \left(1-\frac{\mu}{n}\right)^n \tag{1.6.2}$$

$n \to \infty$の極限では$\left(1-\frac{\mu}{n}\right)^n = \mathrm{e}^{-\mu}$なので（指数関数の定義）

$$\boxed{f(x) \cong \frac{\mu^x}{x!} \mathrm{e}^{-\mu} \equiv P(x,\mu)} \tag{1.6.3}$$

これをポアッソン分布（Poisson distribution）と呼ぶ。

・ポアソン分布の全確率

$$\sum_{x=0}^{\infty} P(x,\mu) = \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = \mathrm{e}^{-\mu}\mathrm{e}^{+\mu} = 1$$

・ポアソン分布の期待値

$$\boxed{E(X) = \sum_{x=0}^{\infty} xP(x,\mu) = \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} x\frac{\mu^x}{x!} = \mu \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = \mu} \tag{1.6.4}$$

期待値μの二項分布の極限をとったこととともコンシステント

・ポアソン分布の分散

$$E(X^2) = \sum_{x=0}^{\infty} x^2 P(x,\mu) = \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} \{x(x-1)+x\} \frac{\mu^x}{x!}$$

$$= \mu^2 \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^{x-2}}{(x-2)!} + \mu \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^{x-1}}{(x-1)!} = \mu^2 \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} + \mu \mathrm{e}^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = \mu^2 + \mu \tag{1.6.5}$$

$$\boxed{V(X) = E(X^2) - \{E(X)\}^2 = \mu^2 + \mu - \mu^2 = \mu} \tag{1.6.6}$$

これは二項分布の分散の極限$np(1-p) \to \mu$ともコンシステント

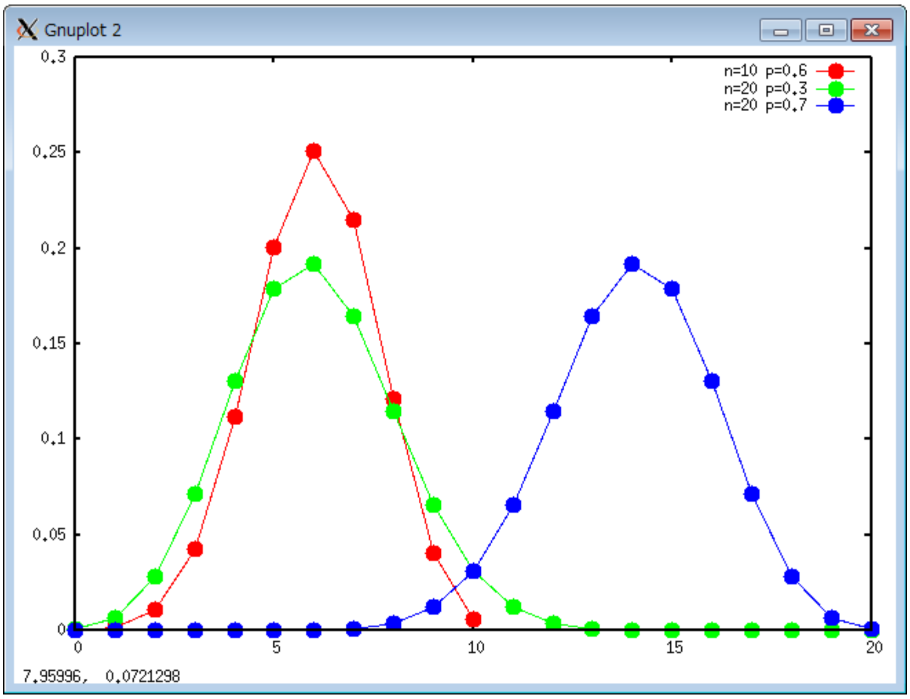


図 1.6.1　n=10,p=0.6(赤)　 n=20,p=0.3(緑)　 n=20,p=0.7(青)の二項分布の確率分布関数。

Bi(10,0.6)と B(20,0.3)の期待値と分散の比較、B(20,0.3)と B(20,0.7)の期待値と分散の比較などを 計算値と比較してみよう。

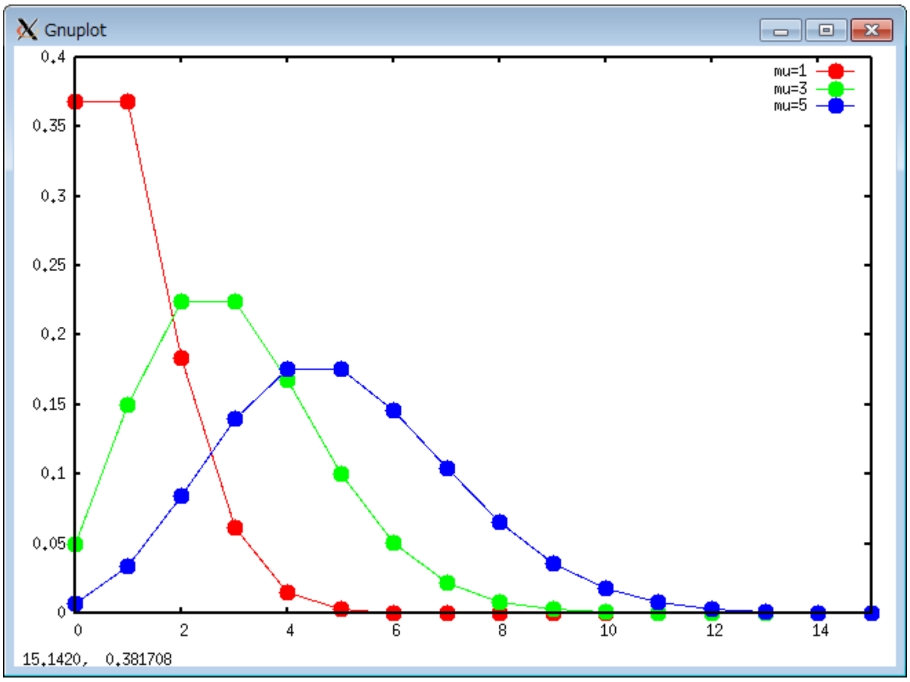


図 1.6.2　μ =1,3,5 のポアソン分布の確率分布関数。

ポアソン分布の理解：

「事象 A が起こる回数の期待値$\mu \equiv np$を固定して$n \to \infty, p \to 0$とした極限をとる」

まれな事象を多数回観測するとポアソン分布になることが多い。自然科学の計測ではよくおこる。

例：放射線計測　10 秒ごとに観測される数少ない計数を数える　*

バクテリアの観測数　ある面積中に観測されるバクテリアの数

1-7 正規分布

・定義

期待値のまわりの相対的なバラツキを表す、二項分布の期待値と標準偏差の比

$$\frac{D(X)}{E(X)} = \frac{\sqrt{np(1-p)}}{np} = \frac{\sqrt{(\frac{1}{p}-1)}}{\sqrt{n}} \tag{1.7.1}$$

は n が大きくなると小さくなる。すなわち、n が大きくなると分布は期待値 np にするどくピークをもつ。ここで、n >>1 の場合を考える。実現値 x >>1 となり x を連続変数として考える。確率分布

$$f(x) = {}_nC_x \cdot p^x \cdot (1-p)^{n-x} \tag{1.7.2}$$

の最大値を求めるため、 $\ln\{f(x)\}$の極値を求めてみよう。

$$\ln\{f(x)\} = \ln\{n!\} - \ln\{x!\} - \ln\{(n-x)!\} + x\ln(p) + (n-x)\ln(1-p) \tag{1.7.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln\{f(x)\} = -\frac{\mathrm{d}}{\mathrm{d}x}\ln\{x!\} - \frac{\mathrm{d}}{\mathrm{d}x}\ln\{(n-x)!\} + \ln(p) - \ln(1-p) \tag{1.7.4}$$

ここで

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln\{x!\} = \lim_{\Delta\to 0}\frac{\ln\{(x+\Delta)!\} - \ln\{x!\}}{\Delta} = \lim_{\Delta\to 0}\frac{1}{\Delta}\ln\left\{\frac{(x+\Delta)!}{x!}\right\} \cong \lim_{\Delta\to 1}\frac{1}{\Delta}\ln\{x+\Delta\}$$
$$= \ln\{x+1\} \cong \ln\{x\} \tag{1.7.5}$$

別解

スターリングの公式

$$\ln\{n!\} \cong n\cdot\ln\{n\} - n \tag{1.7.6}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln\{x!\} = \lim_{\Delta\to 0}\frac{\ln\{(x+\Delta)!\} - \ln\{x!\}}{\Delta} \cong \lim_{\Delta\to 0}\frac{(x+\Delta)\ln\{(x+\Delta)\} - (x+\Delta) - x\ln\{x\} + x}{\Delta}$$
$$= \lim_{\Delta\to 0}\left[\frac{x\ln\left\{\left(1+\frac{\Delta}{x}\right)\right\}}{\Delta} + \ln\{x+\Delta\} - 1\right] = \ln\{x\} \tag{1.7.7}$$

よって

$$\frac{\mathrm{d}}{\mathrm{d}x}\ln\{f(x)\} = -\ln\{x\} + \ln\{n-x\} + \ln(p) - \ln(1-p) = \ln\left(\frac{n-x}{x}\right) - \ln\left(\frac{1-p}{p}\right) \tag{1.7.8}$$

これは、$x = \mu \equiv np$のとき 0 となり、極値をとる。

さらに二階微分をとると

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ln\{f(x)\} = -\frac{1}{x} - \frac{1}{n-x} \tag{1.7.9}$$

$x = \mu \equiv np$のときの値は

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ln\{f(x)\}\right]_{x=\mu} = -\frac{1}{\mu} - \frac{1}{n-\mu} = -\frac{n}{\mu(n-\mu)} = -\frac{1}{np(1-p)} \equiv -\frac{1}{\sigma^2} \quad (1.7.10)$$

さて、$\ln\{f(x)\}$を$x = \mu \equiv np$のまわりでテーラー展開すると

$$\ln\{f(x)\} = \ln\{f(\mu)\} + (x-\mu)\left[\frac{\mathrm{d}}{\mathrm{d}x}\ln\{f(x)\}\right]_{x=\mu} + \frac{(x-\mu)^2}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ln\{f(x)\}\right]_{x=\mu} + \cdots$$

$$\cong \ln\{f(\mu)\} - \frac{(x-\mu)^2}{2\sigma^2} \quad (1.7.11)$$

よって、

$$f(x) = f(\mu)\cdot\exp(-\frac{(x-\mu)^2}{2\sigma^2}) \quad (1.7.12)$$

・正規分布の規格化

ここで、

$$\int_{-\infty}^{+\infty}\exp(-a(x-b)^2)dx = \sqrt{\frac{\pi}{a}} \quad (1.7.13)$$

をつかって、規格化すると

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})} \quad (1.7.14)$$

これを正規分布、またはガウス分布という。

・正規分布の期待値

$$E(X) = \int xf(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int(x-\mu)\cdot\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx + \frac{\mu}{\sqrt{2\pi\sigma^2}}\int\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx$$

$$= \mu \quad (1.7.15)$$

・正規分布の分散

$$\mathrm{V}(X) = \int(x-E(X))^2 f(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int(x-\mu)^2\cdot\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx$$

ここで

$$\int_{-\infty}^{+\infty}\exp(-a(x-b)^2)\,dx = \sqrt{\frac{\pi}{a}}$$

の両辺を a で微分して

$$-\int_{-\infty}^{+\infty}(x-b)^2\exp(-a(x-b)^2)\,dx = -\frac{1}{2a}\sqrt{\frac{\pi}{a}}$$

よって

$$\mathrm{V}(X) = \frac{1}{\sqrt{2\pi\sigma^2}}\int(x-\mu)^2\cdot\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx = \frac{1}{\sqrt{2\pi\sigma^2}}\sigma^2\sqrt{2\pi\sigma^2} = \sigma^2$$

1-8 特性関数

○特性関数

確率変数 X に対して、実数パラメータ ξ の関数として

$$\Phi(\xi)=\langle\exp(i\xi X)\rangle=\int f(x)\exp(i\xi x)dx \tag{1.8.1}$$

を導入して、これを特性関数（characteristic function）とよぶ。（確率関数と同等の情報をもち数学的に扱いやすい。）

指数関数をテーラー展開して

$$\Phi(\xi)=\langle\exp(i\xi X)\rangle=\langle\sum\frac{(i\xi X)^n}{n!}\rangle=\sum\frac{(i\xi)^n}{n!}\langle X^n\rangle \tag{1.8.2}$$

・期待値・分散

特性関数が与えられていると、期待値、分散を求めることができる。

$$\frac{\partial}{\partial\xi}\Phi(\xi)=\int f(x)\frac{\partial}{\partial\xi}\exp(i\xi x)dx=i\int xf(x)\exp(i\xi x)dx \tag{1.8.3}$$

$$\left.\frac{\partial}{\partial\xi}\Phi(\xi)\right|_{\xi=0}=i\int xf(x)dx=i\langle X\rangle \tag{1.8.4}$$

$$\therefore\langle X\rangle=-i\left.\frac{\partial}{\partial\xi}\Phi(\xi)\right|_{\xi=0} \tag{1.8.5}$$

$$\left.\frac{\partial^2}{\partial\xi^2}\Phi(\xi)\right|_{\xi=0}=\int f(x)\left.\frac{\partial^2}{\partial\xi^2}\exp(i\xi x)\right|_{\xi=0}dx=-\int x^2f(x)dx=-\langle X^2\rangle \tag{1.8.6}$$

$$\therefore V(X)=\langle X^2\rangle-\langle X\rangle^2=-\left.\frac{\partial^2}{\partial\xi^2}\Phi(\xi)\right|_{\xi=0}+\left\{\left.\frac{\partial}{\partial\xi}\Phi(\xi)\right|_{\xi=0}\right\}^2 \tag{1.8.7}$$

**

例：正規分布の特性関数

$$\Phi(\xi)=\langle\exp(i\xi X)\rangle=\frac{1}{\sqrt{2\pi\sigma^2}}\int\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\exp(i\xi x)dx$$

$$=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[i\mu\xi-\frac{1}{2}\sigma^2\xi^2\right]\int\exp\left[-\frac{1}{2\sigma^2}(x-\mu-i\sigma^2\xi)^2\right]dx$$

$$=\exp\left[i\mu\xi-\frac{1}{2}\sigma^2\xi^2\right] \tag{1.8.8}$$

期待値と分散を求めてみよう

$$\langle X\rangle=-i\left.\frac{\partial}{\partial\xi}\Phi(\xi)\right|_{\xi=0}=-i\frac{\partial}{\partial\xi}\left.\exp\left[i\mu\xi-\frac{1}{2}\sigma^2\xi^2\right]\right|_{\xi=0}=\mu \tag{1.8.9}$$

$$V(X) = -\frac{\partial^2}{\partial \xi^2}\exp\left[i\mu\xi - \frac{1}{2}\sigma^2\xi^2\right]\Bigg|_{\xi=0} + \left\{\frac{\partial}{\partial \xi}\exp\left[i\mu\xi - \frac{1}{2}\sigma^2\xi^2\right]\Bigg|_{\xi=0}\right\}^2$$

$$= \mu^2 + \sigma^2 - \mu^2 = \sigma^2 \tag{1.8.10}$$

ここで、$\exp\left[(i\xi)\langle X\rangle + \frac{(i\xi)^2}{2}\sigma^2\right]$という量を考えてみよう。ξの 2 次迄の近似で考えると

$$\exp\left[(i\xi)\langle X\rangle + \frac{(i\xi)^2}{2}\sigma^2\right] \cong 1 + (i\xi)\langle X\rangle + \frac{(i\xi)^2}{2}(\sigma^2 + \langle X\rangle^2) + O(\xi^3) = 1 + (i\xi)\langle X\rangle + \frac{(i\xi)^2}{2}\langle X^2\rangle + O(\xi^3)$$

であり、ξの 2 次迄の近似で特性関数と一致する

２：　確率の法則と正規分布

2-1　チェビシェフ不等式

ベルヌーイ試行$B(n,p)$の期待値は$< X >= np$、分散は$\sigma^2 = np(1-p)$である。((1.5.5) (1.5.9))
試行 1 回あたりの事象のおこる割合$Y = {}^{X}\!/_{\mathrm{n}}$を考えると、

期待値は$< Y >= {}^{< X >}\!/_{\mathrm{n}} = p$　(2.1.1)

分散は$V(Y) = {}^{V(X)}\!/_{n^2} = {}^{\sigma^2}\!/_{n^2} = {}^{p(1-p)}\!/_{n}$　(2.1.2)

である。ここで、ベルヌーイ試行$B(n,p)$を行ったときに、$p - \varepsilon < Y = {}^{X}\!/_{\mathrm{n}} < p + \varepsilon$
となる確率を考える。まず、分散の定義から

$$\frac{\sigma^2}{n^2} = \int (y-p)^2 f(y)dy = \int_{p-\varepsilon}^{p+\varepsilon} (y-p)^2 f(y)dy + \int_0^{p-\varepsilon} (y-p)^2 f(y)dy + \int_{p+\varepsilon}^{1} (y-p)^2 f(y)dy \tag{2.1.3}$$

$$\int_{p-\varepsilon}^{p+\varepsilon} (y-p)^2 f(y)dy \geq 0 \tag{2.1.4}$$

$$\int_0^{p-\varepsilon} (y-p)^2 f(y)dy + \int_{p+\varepsilon}^{1} (y-p)^2 f(y)dy \geq \varepsilon^2 \left\{ \int_0^{p-\varepsilon} f(y)dy + \int_{p+\varepsilon}^{1} f(y)dy \right\} = \varepsilon^2 P(|Y-p| \geq \varepsilon) \tag{2.1.5}$$

$$\therefore \frac{\sigma^2}{\varepsilon^2 n^2} \geq P(|Y-p| \geq \varepsilon) \tag{2.1.6}$$

一方$1 = P(|Y-p| \geq \varepsilon) + P(|Y-p| < \varepsilon)$　(2.1.7)
であるので

$$P(|Y-p| < \varepsilon) \leq 1 - \frac{\sigma^2}{\varepsilon^2 n^2} = 1 - \frac{p(1-p)}{\varepsilon^2 n} \tag{2.1.8}$$

どのような ε をとろうとも n を大きく（n>>1/(ε^2)）すると右辺は 1 に限りなく近づく。つまり、試行 1 回あたりの事象のおこる割合は、試行回数をおおきくすると、事象のおこる確率 p に近づいていく。

一般に、確率変数の期待値を$< X >= \mu$、分散はσ^2であるとき、分散の定義から

$$\begin{aligned}\sigma^2 &= \int (x-\mu)^2 f(x)dx = \int_{\mu-\varepsilon}^{\mu+\varepsilon} (x-\mu)^2 f(x)dx + \int_{-\infty}^{\mu-\varepsilon} (x-\mu)^2 f(x)dx + \int_{\mu+\varepsilon}^{+\infty} (x-\mu)^2 f(x)dx \\ &\geq \int_{-\infty}^{\mu-\varepsilon} (x-\mu)^2 f(x)dx + \int_{\mu+\varepsilon}^{+\infty} (x-\mu)^2 f(x)dx \geq \varepsilon^2 \left[\int_{-\infty}^{\mu-\varepsilon} f(x)dx + \int_{\mu+\varepsilon}^{+\infty} f(x)dx \right] \\ &= \varepsilon^2 P(|X-\mu| \geq \varepsilon)\end{aligned} \tag{2.1.9}$$

$$\boxed{\therefore P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}} \tag{2.1.10}$$

これを Chebyshev’s inequality という
確率の和の概念

$$1 = P(|X - \mu| \geq \varepsilon) + P(|X - \mu| < \varepsilon) \tag{2.1.7}$$

から

$$P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2} \tag{2.1.11}$$

となる。これは X が

$$\mu - \varepsilon < X < \mu + \varepsilon \tag{2.1.12}$$

となる確率を表している。

$\varepsilon = \sigma$の時、$P(|X - \mu| < \varepsilon)$ >0　で当たり前、

$\varepsilon = 2\sigma$の時、$P(|X - \mu| < \varepsilon)$>3/4　と導くことができる。

2-2 確率変数の和

N 回の独立な実験(試行)を行ったとき、おのおのの観測値をあらわす確率変数を$X^{(i)}: i = 1 \dots N$とする。また、それぞれの期待値と分散を

$$< X^{(i)} \geq \int x_i f^{(i)}(x_i) dx_i = \mu_i \qquad (2.2.1)$$

$$< (X^{(i)} - \mu_i)^2 >= {\sigma_i}^2 : i = 1 \dots N \qquad (2.2.2)$$

とする。

ここで、N 回の観測値の和を考える

$$X(\mathrm{N}) \equiv \sum_i X^{(i)} \qquad (2.2.3)$$

$X(\mathrm{N})$の期待値は、

$$< X(\mathrm{N}) >= \sum_i < X^{(i)} > = \sum_i \mu_i \qquad (2.2.4)$$

分散は

$$\langle (X(\mathrm{N}) - < X(\mathrm{N}) >)^2 \rangle = \langle \left(\sum_i \left(X^{(i)} - \mu_i \right) \right)^2 \rangle = \langle \left(\sum_i \left(X^{(i)} - \mu_i \right) \right) \left(\sum_j \left(X^{(j)} - \mu_j \right) \right) \rangle$$

$$= \langle \sum_{i \neq j} \left(X^{(i)} - \mu_i \right) \left(X^{(j)} - \mu_j \right) \rangle + \langle \sum_i \left(X^{(i)} - \mu_i \right)^2 \rangle \qquad (2.2.5)$$

また、$P(X^{(i)} = x_i, X^{(j)} = x_j) = f^{(i)}(x_i) f^{(j)}(x_j)$ (2.2.6)

であるので、

$$\langle \sum_{i \neq j} \left(X^{(i)} - \mu_i \right) \left(X^{(j)} - \mu_j \right) \rangle = \iint \sum_{i \neq j} (x - \mu_i)(y - \mu_j) f^{(i)}(x) f^{(j)}(y) dx dy$$

$$= \sum_{i \neq j} \int (x - \mu_i) f^{(i)}(x) dx \int (y - \mu_j) f^{(j)}(y) dx = 0 \qquad (2.2.7)$$

$$\therefore < (X(\mathrm{N}) - < X(\mathrm{N}) >)^2 >= \sum {\sigma_\mathrm{i}}^2 \qquad (2.2.8)$$

<<第 3 講ここまで>>

<<第 4 講>>

2-3 大数の法則

N 回の独立な実験(試行)を行い、おのおの確率変数$X^{(i)}: i = 1 \ldots N$が同じ確率分布(正規分布でなくて構わない)に従うとする。このときの期待値と分散を

$$< X^{(i)} >= \mu \tag{2.3.1}$$

$$< (X^{(i)} - \mu)^2 >= \sigma^2 \tag{2.3.2}$$

とする。ここで、N 回の観測値の平均

$$\bar{X} \equiv \frac{X(\mathrm{N})}{\mathrm{N}} = \frac{1}{\mathrm{N}} \sum X^{(i)} \tag{2.3.3}$$

を考えると、期待値と分散は

$$\langle \bar{X} \rangle = \frac{1}{\mathrm{N}} \sum \langle X^{(i)} \rangle = \mu \tag{2.3.4}$$

$$\langle (\bar{X} - \mu)^2 \rangle = \frac{1}{\mathrm{N}^2} \sum \sigma^2 = \frac{\sigma^2}{\mathrm{N}} \tag{2.3.5}$$

ここで、チェビシェフの不等式をつかうと、

$$\boxed{P(|\bar{X} - \mu| < \varepsilon) \leq 1 - \frac{\sigma^2}{N\varepsilon^2}} \tag{2.3.6}$$

つまり、N→∞のとき、$\bar{X}$は限りなくμに近づく。測定回数を多くするほど真の値に近い結果を得ることができる。

2-4 中心極限定理

N 回の独立な実験(試行)を行い、おのおの確率変数$X^{(i)}: i = 1 \dots N$が同じ確率分布(正規分布でなくて構わない)に従うとする。このときの期待値と分散を

$$< X^{(i)} >= \mu \qquad (2.3.1)$$

$$< (X^{(i)} - \mu)^2 >= \sigma^2 \qquad (2.3.2)$$

とする。ここで、標本平均の確率分布$\bar{X}$は N が大きくなるにつれて期待値μ、分散$\frac{\sigma^2}{N}$の正規分布に近づく。

$$\bar{Z} \equiv \frac{\bar{X}-\mu}{\sigma/\sqrt{N}} \qquad (2.4.1)$$

を定義すると、期待値 0 、分散 1 になる。((2.3.4)、(2.3.5)より。)

また、確率変数$X^{(i)}$に対して、

$$Z^{(i)} \equiv \frac{X^{(i)}-\mu}{\sigma} \qquad (2.4.2)$$

を定義すると、これも期待値 0 、分散 1 になり、$\bar{Z}$は

$$\bar{Z} = \frac{1}{\sqrt{N}}\sum Z^{(i)} \qquad (2.4.3)$$

となる。

ここで、$\bar{Z}$に対する特性関数を考えると

$$\Phi(i\xi) = \langle \exp(i\xi\bar{Z}) \rangle = \langle \prod \exp(i\xi \frac{Z^{(i)}}{\sqrt{N}}) \rangle = \prod \langle \exp(i\xi \frac{Z^{(i)}}{\sqrt{N}}) \rangle = \prod \Phi^{(i)}(i\xi) = \left(\Phi^{(i)}(i\xi)\right)^N \qquad (2.4.4)$$

ここで$\Phi^{(i)}(i\xi)$は、$Z^{(i)}/\sqrt{N}$に対する特性関数であるので、ξの 2 次迄の近似で

$$\Phi^{(i)}(\xi) = exp\left[\frac{i\xi}{\sqrt{N}}\langle Z^{(i)} \rangle + \frac{1}{2}\left(\frac{i\xi}{\sqrt{N}}\right)^2 V\left(Z^{(i)}\right) + O(\left(\frac{i\xi}{\sqrt{N}}\right)^3)\right] \qquad (2.4.5)$$

と展開できる。$Z^{(i)}$の期待値 0 、分散 1 であるので、

$$\langle \left(Z^{(i)}\right)^1 \rangle = 0, \langle \left(Z^{(i)}\right)^2 \rangle = 1 \qquad (2.4.6)$$

である。従って、

$$\Phi^{(i)}(i\xi) = exp\left[\frac{1}{2}\left(\frac{i\xi}{\sqrt{N}}\right)^2 + O(\left(\frac{i\xi}{\sqrt{N}}\right)^3)\right] \qquad (2.4.7)$$

よって、

$$\Phi(i\xi) = exp\left[\frac{1}{2}(i\xi)^2 + N \times O(\left(\frac{i\xi}{\sqrt{N}}\right)^3)\right] \qquad (2.4.8)$$

N→∞の極限をとると、

$$\Phi(i\xi) \to exp\left[\frac{1}{2}(i\xi)^2\right] \tag{2.4.9}$$

これは、期待値０、分散１の正規分布の特性関数と一致する
すなわち、

同じ確率分布に従う N ケの独立な確率変数$X^{(i)}: i = 1 \ldots N$から

$$\bar{Z} \equiv \frac{\bar{X}-\mu}{\sigma/\sqrt{\mathrm{N}}} \tag{2.4.10}$$

をつくると、$\bar{Z}$の分布は N→∞の極限で期待値０、分散１の正規分布となる

2-5 正規分布の性質

○確率変数の和

独立な確率変数$X^{(i)}: i = 1 \ldots N$は、（それぞれ別の）正規分布に従うとする。これらの確率変数の和

$$X(\mathrm{N}) \equiv \sum_i X^{(i)} \tag{2.5.1}$$

を考える。

$X(\mathrm{N})$の特性関数は、

$$\Phi(\xi) = \langle \exp\left(i\xi X(\mathrm{N})\right)\rangle = \langle \exp\left(i\xi \sum_i X^{(i)}\right)\rangle = \iint .. \int \exp(i\xi \sum_i x_i)\, f(x_1, x_2, .. x_n)\, dx_1 dx_2 .. dx_n$$
(2.5.2)

確率変数は独立で、$f(x_1, x_2, .. x_n) = \prod f_i(x_i)$であるから、

$$\Phi(\xi) = \prod \int \exp(i\xi x_i)\, f_i(x_i)\, dx_i = \prod \langle \exp\left(i\xi X^{(i)}\right)\rangle = \prod \exp\left[i\xi\mu_i + (i\xi)^2 \frac{\mathrm{\sigma_i}^2}{2}\right]$$

$$= \exp\left[i\xi \sum \mu_i + (i\xi)^2 \sum \frac{\mathrm{\sigma_i}^2}{2}\right] \tag{2.5.3}$$

これから、$X(\mathrm{N})$も正規分布であり、

$$\langle X(\mathrm{N})\rangle = \frac{\partial}{\partial(i\xi)} \Phi(\xi)\bigg|_{\xi=0} = \sum \mu_i \tag{2.5.4}$$

$$\langle (X(\mathrm{N}))^2\rangle = \frac{\partial^2}{\partial(i\xi)^2} \Phi(\xi)\bigg|_{\xi=0} = \left(\sum \mu_i\right)^2 + \sum \mathrm{\sigma_i}^2 \tag{2.5.5}$$

$$V(X(\mathrm{N})) = \langle (X(\mathrm{N}))^2\rangle - (\langle X(\mathrm{N})\rangle)^2 = \sum \mathrm{\sigma_i}^2 \tag{2.5.6}$$

であることが判る。

○確率変数の平均

独立な確率変数$X^{(i)}: i = 1 \ldots N$は、同一の正規分布に従うとする。これらの確率変数の平均

$$\bar{X} \equiv \frac{1}{\mathrm{N}} \sum_i X^{(i)} \tag{2.5.7}$$

を考える。特性関数は

$$\Phi_{\bar{X}}(\xi) = \exp\left[i\xi \sum \frac{\mu_i}{N} + (i\xi)^2 \sum \frac{\mathrm{\sigma_i}^2}{2\mathrm{N}^2}\right] = \exp\left[i\xi\mu + (i\xi)^2 \frac{\sigma^2}{2\mathrm{N}}\right] \tag{2.5.8}$$

となり、

$$\langle \bar{X}\rangle = \mu \tag{2.5.9}$$

$$V(\bar{X}) = \frac{\sigma^2}{N} \tag{2.5.10}$$

の正規分布となる。

*

3　標本・母集団・推定

3-1 標本と母集団

母集団（population）とは、統計学的推定で基本として仮定する、ある要素の集合であって、これからランダムな標本を抽出して観察し、その結果から逆に母集団を推定するという形で用いる。

標本（sample）とは、母集団の部分集合のこと。この母集団から部分集合を対応させる規則（可測関数）を標本抽出（sampling）と呼ぶ、

母集団を表現する数値を母数（またはパラメータ）というのに対し、標本を表現する数値を統計量(statistics)という。統計量は標本から算出される数値である。また統計量で特に母数を推定するために用いられるものを推定量という。

母集団を完全に調査するのは不可能な場合、標本から母集団の特性を推定する必要がある。この標本抽出には作為抽出法と無作為抽出法（random sampling）の２つの抽出方法があり、統計学では無作為抽出法だけを議論する。

無作為抽出法：母集団のどの要素も等しい確率で標本に選ばれる。

また統計学的推定（標本から母集団の性質を推定する）を行うには、各標本に対し、それが選抜される確率を知る必要がある。このために様々な標本抽出法が開発されており、例えば異なる標本が選抜される確率がすべて同じならば、その抽出法は単純ランダム（無作為）抽出という。

母集団から N ケの標本を抽出する

$$x^{(i)}: i = 1 \ldots N \qquad (3.1.1)$$

標本を特徴づける標本平均値(sample mean value)を導入する

$$\bar{x} = \frac{1}{N}\sum x^{(i)} \qquad (3.1.2)$$

次に、標本平均値のまわりのバラツキをあらわす標本分散値(sample variance value)を導入する

$$v^2 = \frac{1}{N}\sum\left(x^{(i)} - \bar{x}\right)^2 \qquad (3.1.3)$$

この標本分散値の平方

$$v = \sqrt{v^2} \qquad (3.1.4)$$

を、標本標準偏差という

3-2 標本確率変数

N ケの標本$x^{(i)}: i = 1 \dots N$を独立な確率変数$X^{(i)}: i = 1 \dots N$の実現値だと考える。標本は同じ母集団から抽出されたのだから、これら N ケの確率変数は同一の確率分布に従う。
すなわち、
「大きな母集団から$x^{(i)}: i = 1 \dots N$という標本を抽出することは、母集団と同じ確率的な構造を持った互いに独立な確率変数$X^{(i)}: i = 1 \dots N$の実現値が$x^{(i)}: i = 1 \dots N$で与えられた事と等価である」
また、互いに独立で同一の確率分布に従うものを、i.i.d.(independent identically distributed)な確率変数という。

これらの標本確率変数から統計量として、標本平均(sample mean)

$$\bar{X} = \frac{1}{\mathrm{N}} \sum_i X^{(i)} \tag{3.2.1}$$

及び標本分散(sample variance)

$$V^2 = \frac{1}{N} \sum \left(X^{(i)} - \bar{X}\right)^2 \tag{3.2.2}$$

標本標準偏差

$$V = \sqrt{V^2} \tag{3.2.3}$$

を導入する.(これらも確率変数である)

注目している確率変数の母集団に関する期待値を母平均μ、その分散を母分散σ ²という。
○標本平均及び標本分散と、母平均、母分散との関係

$$\boxed{\langle \bar{X} \rangle = \frac{1}{\mathrm{N}} \sum_i \langle X^{(i)} \rangle = \frac{1}{\mathrm{N}} \sum_i \mu = \mu} \tag{3.2.4}$$

$$\begin{aligned} \langle V^2 \rangle &= \frac{1}{N} \sum \langle (X^{(i)} - X)^2 \rangle \\ &= \frac{1}{N} \sum \langle \left\{ \left(X^{(i)} - \mu\right) - (\bar{X} - \mu) \right\}^2 \rangle \\ &= \frac{1}{N} \sum \langle \left(X^{(i)} - \mu\right)^2 \rangle + \frac{1}{N} \sum \langle (\bar{X} - \mu)^2 \rangle - \frac{2}{N} \sum \langle \left(X^{(i)} - \mu\right)(\bar{X} - \mu) \rangle \end{aligned} \tag{3.2.5}$$

$$\frac{1}{N} \sum \langle \left(X^{(i)} - \mu\right)^2 \rangle = \frac{1}{N} \sum \sigma^2 = \sigma^2 \tag{3.2.6}$$

$$\frac{1}{N} \sum \langle (\bar{X} - \mu)^2 \rangle = \langle (\bar{X} - \mu)^2 \rangle = \frac{\sigma^2}{N} \tag{3.2.7}$$

$$\frac{2}{N} \sum \langle \left(X^{(i)} - \mu\right)(\bar{X} - \mu) \rangle = 2 \langle \left\{ \frac{1}{N} \sum \left(X^{(i)} - \mu\right) \right\} (\bar{X} - \mu) \rangle = 2 \langle (\bar{X} - \mu)^2 \rangle = 2 \frac{\sigma^2}{N} \tag{3.2.7}$$

$$\boxed{\therefore \langle V^2 \rangle = \frac{N-1}{N} \sigma^2} \tag{3.2.8}$$

そこで不偏分散(unbiased variance)

$$\boxed{S^2 \equiv \frac{N}{N-1} V^2 = \frac{1}{N-1} \sum \left(X^{(i)} - \bar{X} \right)^2} \qquad (3.2.9)$$

を導入すると、

$$\langle S^2 \rangle = \sigma^2 \qquad (3.2.10)$$

となる。 **

3-3 推定

○統計的推定

標本抽出によって得られたデータを基に、母集団の情報を得る操作を統計的推計（statistical estimation）という。

このとき、母集団分布を仮定し、分布を特徴づける母数（パラメータ）をデータから決定するやり方を、パラメトリックモデルによる推定法という。

例えば、母集団を正規分布だと仮定したときに、分布を特徴づけるパラメータは平均μと分散σ^2である。また、母集団をポアッソン分布だと仮定したときのパラメータは平均μである。

○点推定

母数をθとする。抽出されたデータを用いてθを決定する手続きを点推定（point estimation）という。まず、母数θに対応する確率変数Θをさがす。例えば、平均μという母数に対応する確率変数としては標本平均がΘとなる。このΘは確率変数$X^{(i)}: i = 1 \dots N$の関数であり、

$\Theta = \Theta(X^{(i)})$

Θを統計推定量（statistical estimator）又は推定量（estimator）と呼ぶ。特に

$\langle\Theta\rangle = \theta$

となるΘを不偏推定量（unbiased estimator）と呼ぶ。

そこで、標本から計算される不偏推定量の実現値を、母数として推定する。

例：母集団分布が正規分布だと仮定した場合、

標本平均　$\bar{X} = \frac{1}{\mathrm{N}}\sum_i X^{(i)}$　(3.3.1)

不偏分散　$S^2 = \frac{1}{N-1}\sum\left(X^{(i)} - \bar{X}\right)^2$　(3.3.2)

が、それぞれ平均μと分散σ^2の不偏推定量である。

そこで、母集団から抽出された N ケの標本を$x^{(i)}: i = 1 \dots N$から計算される

標本平均値　$\bar{x} = \frac{1}{N}\sum x^{(i)}$　(3. 3. 3)

標本分散値　$v^2 = \frac{1}{N}\sum\left(x^{(i)} - \bar{x}\right)^2$　(3.3.4)

を、それぞれ母平均μと母分散σ^2として推定する。

標本平均値が母集団の平均と等しいわけではない、推定しただけである。ただし、大数の法則から、標本数Nが大きければ、標本平均値が母集団の平均に限りなく近づくことが言える。

<<第 4 講ここまで>>

<<第 5 講>>

3-4 区間推定

点推定で推定したパラメータのバラツキや信頼区間を示すことを区間推定(interval estimation)という。従ってαを 0 以上 1 以下の数として、

$P(\theta_l \le \theta \le \theta_u) = 1 - \alpha$

となるように決める。つまり「$1-\alpha$の確かさで、母数は$\theta_l \le \theta \le \theta_u$の範囲にある」という形式になる。ここで、$1-\alpha$を信頼水準(confidence level)、$\theta_l, \theta_u$を信頼限界（confidence limit）と呼ぶ。
N ケの標本$x^{(i)}: i = 1 \dots N$を独立な確率変数$X^{(i)}: i = 1 \dots N$の実現値である。統計推定量$\Theta = \Theta(X^{(i)})$の分布を母数 θ と共に考える。

○母平均の μ の区間推定

母集団は正規分布$N(\mu, \sigma^2)$であり、母分散 σ ^{2}が既知であるとする。

この時、母平均 μ に対する推定量は標本平均　$\bar{X} = \frac{1}{N}\sum_i X^{(i)}$であり、

標本平均は正規分布$N(\mu, \sigma^2/N)$に従う。(中心極限定理)

そこで、

$$\bar{Z} \equiv \frac{\bar{X}-\mu}{\sigma/\sqrt{N}} \tag{3.4.1}$$

を考えると、これは正規分布$N(0,1)$に従うのであるから、

$$P(z_l \le \bar{Z} \le z_u) = \int_{z_l}^{z_u} \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) dz = 1 - \alpha \tag{3.4.2}$$

となるように、z_l, z_uを決定する。正規分布$N(0,1)$に対しては、$z_l = -z_u$であるから、

$$P\left(\bar{X} - \frac{\sigma}{\sqrt{N}} z_u \le \mu \le \bar{X} + \frac{\sigma}{\sqrt{N}} z_u\right) = 1 - \alpha \tag{3.4.3}$$

である。
一般には、99%の信頼度($1-\alpha = 0.99$)、95%の信頼度($1-\alpha = 0.95$)や 90%の信頼度($1-\alpha = 0.90$)がよく使われる。この時の信頼限界の値は

99%の信頼度（$1-\alpha = 0.99$）	$z_u = 2.58$	(3.4.5)
95%の信頼度（$1-\alpha = 0.95$）	$z_u = 1.96$	(3.4.6)
90%の信頼度（$1-\alpha = 0.90$）	$z_u = 1.65$	(3.4.7)
99.7%の信頼度（$1-\alpha = 0.997$）	$z_u = 3.$	(3.4.8)
95%の信頼度（$1-\alpha = 0.90$）	$z_u = 2.$	(3.4.9)
68%の信頼度（$1-\alpha = 0.68$）	$z_u = 1.$	(3.4.10)

となる。(付表 1 参照)

○母分散 σ

母分散 σ ^{2}が既知であるとして、母平均 μ に対する推定を行ったわけだが、母分散の仮定が正しい

ことを検証するにはどうすればよいであろうか。それには標本分散の分布を用いる。

$$\mathrm{S} = \frac{\mathrm{N}}{\sigma^2} v^2 = \frac{1}{\sigma^2} \sum \left(x^{(i)} - \bar{x}\right)^2 \qquad (3.4.11)$$

は、自由度 $N-1$ の χ^2 分布に従うことが判っている
ここで、自由度 N の χ^2 分布の確率密度関数は

$$\mathrm{C}(y, N) = \left(\frac{1}{2}\right)^{\frac{N}{2}} \frac{1}{\Gamma\left(\frac{N}{2}\right)} y^{\frac{N}{2}-1} \exp\left(-\frac{y}{2}\right) \qquad (3.4.12)$$

で与えられる。
自由度 N の χ^2 分布に従う確率変数 Y の期待値と分散は

$$\langle Y \rangle = \int_0^\infty y\mathrm{C}(y, N) dy = \frac{2}{\Gamma\left(\frac{N}{2}\right)} \int_0^\infty x^{\frac{N}{2}} \exp(-x) dx = \frac{2\Gamma\left(\frac{N}{2}+1\right)}{\Gamma\left(\frac{N}{2}\right)} = N \qquad (3.4.13)$$

$$\langle Y^2 \rangle = \int_0^\infty y^2\mathrm{C}(y, N) dy = \frac{4}{\Gamma\left(\frac{N}{2}\right)} \int_0^\infty x^{\frac{N}{2}+1} \exp(-x) dx = \frac{4\Gamma\left(\frac{N}{2}+2\right)}{\Gamma\left(\frac{N}{2}\right)} = N(N+2) \qquad (3.4.14)$$

$$\langle Y^2 \rangle - \langle Y \rangle^2 = 2N \qquad (3.4.15)$$

そこで、reduced χ^2

$$\tilde{S} \equiv \frac{S}{N-1} \qquad (3.4.16)$$

を定義すると、

$$\langle \tilde{\mathrm{S}} \rangle = 1 \qquad (3.4.17)$$

$$\langle \tilde{S}^2 \rangle - \langle \tilde{S} \rangle^2 = \frac{2}{N-1} \qquad (3.4.18)$$

よって、reduced χ^2 が 1 に近いことを確認すればよい。
このとき、信頼度 α での reduced χ^2 の値、つまり

$$\int_{\mathrm{S}_1}^\infty \mathrm{C}(y\mathrm{N}, N)\, dy = 1 - \frac{\alpha}{2} \qquad (3.4.19)$$

$$\int_0^{\mathrm{S}_2} \mathrm{C}(y\mathrm{N}, N)\, dy = 1 - \frac{\alpha}{2} \qquad (3.4.20)$$

となる S_1, S_2 の値は下表のとおりである

自由度　N	信頼度　α=90%		信頼度　α=99%	
	S_1	S_2	S_1	S_2
1	0	3.84	0	7.88
5	0.23	2.20	0.08	3.35
10	0.39	1.83	0.22	2.52
50	0.69	1.44	0.54	1.59
100	0.78	1.24	0.67	1.40

〇母平均の μ の区間推定（母分散 σ^2 が未知の場合）

この時も、母平均 μ に対する推定量は標本平均　$\bar{X}=\frac{1}{\mathrm{N}}\sum_i X^{(i)}$であり、

標本平均は正規分布$N(\mu,{}^{\sigma^2}/_{\mathrm{N}})$に従う。

しかし、母分散 σ^2 が未知なので、$\bar{Z}$が定義できない。

そこで、母分散 σ^2 の不偏推定量である不偏分散　$S^2=\frac{1}{N-1}\sum\left(X^{(i)}-\bar{X}\right)^2$を使って

$$T\equiv\frac{\bar{X}-\mu}{{}^{\mathrm{S}}/_{\sqrt{\mathrm{N}}}} \tag{3.4.21}$$

を考えてみると、これは自由度 N－1 の t-分布（スチューデント分布）に従うことが判っている。
ここで、自由度 N－1 の t-分布の確率密度関数は

$$T(t,N-1)=\frac{\Gamma\left(\frac{N}{2}\right)}{\sqrt{\pi(N-1)}\Gamma\left(\frac{N-1}{2}\right)}\frac{1}{\left(\frac{t^2}{N}\right)^{\frac{N}{2}}} \tag{3.4.22}$$

で与えられる。(N→∞で正規分布)
ここでΓ関数は

$$\Gamma(\lambda)=\int_0^\infty x^{\lambda-1}\exp(-x)dx \tag{3.4.23}$$

定義され、

$$\Gamma(\lambda+1)=\lambda\Gamma(\lambda)$$

$$\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi} \tag{3.4.24}$$

であることより

$$\Gamma\left(\frac{N}{2}\right)=\left[\begin{array}{ll}(n-1)! & N=2n\\ \sqrt{\pi}\left(\frac{1}{2}\right)^n\prod_{i=0}^{n-1}(2i+1) & N=2n+1\end{array}\right.\quad n=1,2,3\ldots \tag{3.4.25}$$

信頼限界の値は、自由度 Nに依存し

	N=10	N=100
99%の信頼度（$1-\alpha=0.99$）	$z_u=3.25$	$z_u=2.63$
95%の信頼度（$1-\alpha=0.95$）	$z_u=2.26$	$z_u=1.98$
90%の信頼度（$1-\alpha=0.90$）	$z_u=1.83$	$z_u=1.66$

となる。

3-5　Null-result の観測：Poisson 推定

母集団が平均 μ<1 の Poisson 分布に従っているとする。

○測定で 0 ケの事象が観測されたとき、推定によって μ の上限値を見積もる。
従って α を 0 以上 1 以下の数として、
$P(\mu \leq \mu_u) = 1 - \alpha$
となるように決める。

$$\frac{\mu^x}{x!}\mathrm{e}^{-\mu} \equiv P(x,\mu) \tag{1.6.3}$$

より　$P(0,\mu) = \mathrm{e}^{-\mu}$　(3.5.1)
であるから

$$P(\mu \leq \mu_u) = \int_0^{\mu_u} \exp(-y)dy = 1 - \exp(-\mu_u) = 1 - \alpha \tag{3.5.2}$$

となるようにすればよい。つまり、
$\mu_u = -\ln(\alpha)$
信頼度 90%の上限値(1- α =0.9)　$\mu_u = \ln(10) = 2.30$
信頼度 95%の上限値(1- α =0.95)　$\mu_u = \ln(20) = 2.99$
信頼度 99%の上限値(1- α =0.99)　$\mu_u = \ln(100) = 4.60$

○同様に測定で 1 ケの事象が観測されたとき、推定によって μ の上限値を見積もる。(1.6.3)より

$$P(1,\mu) = \mu\mathrm{e}^{-\mu} \tag{3.5.3}$$

であるから

$$P(\mu \leq \mu_u) = \int_0^{\mu_u} y\exp(-y)dy = 1 - (1 + \mu_u)\exp(-\mu_u) = 1 - \alpha \tag{3.5.4}$$

となるようにすればよい。つまり、
信頼度 90%の上限値(1- α =0.9)　$\mu_u \approx 3.8$
信頼度 95%の上限値(1- α =0.95)　$\mu_u \approx 5.8$
信頼度 99%の上限値(1- α =0.99)　$\mu_u \approx 6.6$

その他の観測値に対しての上限値もしくは下限値（片側のみに α を考える）は付表 2 に示す。観測値に対しての上限値と下限値を（両側を合わせて α を考える）場合は付表 3 に示す。

3-6 最尤法

目的は引き続き「母数の推定」

母集団の確率密度関数を母数θに依存することを明示して$W(x) = W(x, \theta)$と書く。
N ケの標本$x^{(i)}: i = 1 \ldots N$を独立な確率変数$X^{(i)}: i = 1 \ldots N$の実現値だから、母数θに対して N ケの標本値$x^{(i)}: i = 1 \ldots N$が実現する確率を

$$L(\theta) \equiv W\left(x^{(1)}, \theta\right) \mathrm{W}\left(x^{(2)}, \theta\right) \ldots \mathrm{W}\left(x^{(N)}, \theta\right) = \prod W\left(x^{(i)}, \theta\right) \tag{3.6.1}$$

を尤度関数（likelihood function）として定義する。
確率の高い状態ほど起こりやすいのだから、$L(\theta)$が最大となるように母数θを決めればよい。これを最尤法（maximum likelihood method）といい、このようにして決めた値を$\hat{\theta}$とかいて最尤推定値（maximum likelihood estimate）と呼ぶ。

実際の手順としては、尤度関数の対数(対数尤度)

$$l(\theta) \equiv \ln(L(\theta)) = \sum \ln\left(W\left(x^{(i)}, \theta\right)\right) \tag{3.6.2}$$

を最大になるように、つまり

$$\left.\frac{d}{d\theta} l(\theta)\right|_{\theta=\hat{\theta}} = 0 \tag{3.6.3}$$

となるように$\hat{\theta}$を決定する。

例：母集団分布が正規分布だと仮定した場合、
母数は平均μと分散σ^2であるから、確率密度関数は

$$W(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{3.6.4}$$

対数尤度は

$$l(\mu, \sigma^2) = \sum \ln\left(W\left(x^{(i)}, \mu, \sigma^2\right)\right) = -\frac{1}{2\sigma^2} \sum \left(x^{(i)} - \mu\right)^2 - \frac{\mathrm{N}}{2} \ln(2\pi\sigma^2) \tag{3.6.5}$$

最大になる条件は

$$0 = \frac{\partial}{\partial\mu} l(\mu, \sigma^2) = -\frac{1}{\sigma^2} \sum \left(x^{(i)} - \mu\right) \tag{3.6.6}$$

$$0 = \frac{\partial}{\partial(\sigma^2)} l(\mu, \sigma^2) = \frac{1}{2(\sigma^2)^2} \sum \left(x^{(i)} - \mu\right)^2 - \frac{\mathrm{N}}{2\sigma^2} \tag{3.6.7}$$

よって、

$$\hat{\mu} = \frac{1}{N} \sum x^{(i)} = \bar{x} \tag{3.6.8}$$

$$\hat{\sigma}^2 = \frac{1}{N} \sum \left(x^{(i)} - \mu\right)^2 = v^2 \tag{3.6.9}$$

となる。

**

3-7 最小二乗法

物理現象の測定データには、誤差が含まれ、それは系統誤差と偶然誤差を含んでいる。このうち、偶然誤差は、測定における信号経路の微視的現象に由来するならば、正規分布になる(白色雑音である)と期待されることが多い。測定データがモデル関数と白色雑音の和で表わせるならば、測定データとモデル関数の差の分散を最小にするようにモデル関数を決定することができる。

パラメータ$t = t^{(i)}: i = 1 \ldots N$で規定される N ケの測定点での測定値を$x^{(i)}: i = 1 \ldots N$とする。測定値が、未知のパラメータ θ を含むモデル関数$f(t,\theta)$と、正規分布に従う誤差(分散${\sigma_i}^2: i = 1 \ldots N$)の和であるとする。すると、$x^{(i)}: i = 1 \ldots N$は、正規分布$N\left(f\left(t^{(i)},\theta\right),{\sigma_i}^2\right)$に従う独立な確率変数$X^{(i)}: i = 1 \ldots N$の実現値と考えることができ、最尤法を適応してパラメータ θ を決定する。

確率密度関数は

$$W(x,t,\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-f(t,\theta))^2}{2\sigma^2}\right) \tag{3.7.1}$$

対数尤度を最大にする条件は

$$0 = \frac{\partial}{\partial\theta} l(\theta) = \frac{\partial}{\partial\theta}\sum \ln\left(W\left(x^{(i)},t^{(i)},\theta,{\sigma_i}^2\right)\right) = -\frac{\partial}{\partial\theta}\sum\frac{1}{2{\sigma_i}^2}\left(x^{(i)} - f\left(t^{(i)},\theta\right)\right)^2 \tag{3.7.2}$$

つまり

$$\boxed{S = \sum\frac{1}{{\sigma_i}^2}\left(x^{(i)} - f\left(t^{(i)},\theta\right)\right)^2 >} 0 \tag{3.7.3}$$

を最小にするようにパラメータ θ を決定する。

例： モデル関数を一次関数$f(t,a,b) = at + b$とする。

$$S(a,b) = \sum\frac{1}{{\sigma_i}^2}\left(x^{(i)} - at^{(i)} - b\right)^2 \tag{3.7.4}$$

最小になる条件は

$$0 = \frac{\partial}{\partial a}S(a,b) = \sum\frac{1}{{\sigma_i}^2}\left(x^{(i)} - at^{(i)} - b\right)t^{(i)} \tag{3.7.5}$$

$$0 = \frac{\partial}{\partial b}S(a,b) = \sum\frac{1}{{\sigma_i}^2}\left(x^{(i)} - at^{(i)} - b\right) \tag{3.7.6}$$

よって、

$$\begin{pmatrix} \sum\frac{x^{(i)}t^{(i)}}{{\sigma_i}^2} \\ \sum\frac{x^{(i)}}{{\sigma_i}^2} \end{pmatrix} = \begin{pmatrix} \sum\frac{\left(t^{(i)}\right)^2}{{\sigma_i}^2} & \sum\frac{t^{(i)}}{{\sigma_i}^2} \\ \sum\frac{t^{(i)}}{{\sigma_i}^2} & \sum\frac{1}{{\sigma_i}^2} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \tag{3.7.7}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{D}\begin{pmatrix} \sum\frac{1}{{\sigma_i}^2} & -\sum\frac{t^{(i)}}{{\sigma_i}^2} \\ -\sum\frac{t^{(i)}}{{\sigma_i}^2} & \sum\frac{\left(t^{(i)}\right)^2}{{\sigma_i}^2} \end{pmatrix}\begin{pmatrix} \sum x^{(i)}\,t^{(i)} \\ \sum x^{(i)} \end{pmatrix} \tag{3.7.8}$$

$$D = \left(\sum \frac{\left(t^{(i)}\right)^2}{\sigma_i{}^2}\right)\left(\sum \frac{1}{\sigma_i{}^2}\right) - \left(\sum \frac{\left(t^{(i)}\right)}{\sigma_i{}^2}\right)^2 \tag{3.7.9}$$

N ケの測定点での測定値を
$x^{(i)}: i = 1 \ldots N$, 誤差分散$\sigma_{xi}{}^2: i = 1 \ldots N$）
及び
$y^{(i)}: i = 1 \ldots N$, 誤差分散$\sigma_{yi}{}^2: i = 1 \ldots N$）
とする。
測定値の組み$(x \quad y)$の間に、未知のパラメータ θ を含むモデル関数$f(x, \theta)$を使って$y = f(x, \theta)$という関係が成り立つものとする。
このとき、最小二乗法を使ってパラメータ θ を決定するには、
誤差の伝搬の式を使って、先の例で

$$\sigma_i{}^2 \to \sigma_{yi}{}^2 + \left(\left.\frac{\partial f(x,\theta)}{\partial x}\right|_{x=x_i}\right)^2 \sigma_{xi}{}^2 \tag{3.7.10}$$

の置き換えをすればよい。

いま導入した

$$S = \sum \frac{1}{\sigma_i{}^2}\left(x^{(i)} - f\left(t^{(i)}, \theta\right)\right)^2 \tag{3.7.11}$$

は、パラメータ θ の個数を m としたとき、自由度 $N-m$ の χ ^{2}分布に従うことが判っている

ここで、自由度 N の χ ^{2}分布の確率密度関数は

$$\mathrm{C}(y, N) = \left(\frac{1}{2}\right)^{\frac{N}{2}} \frac{1}{\Gamma\left(\frac{N}{2}\right)} y^{\frac{N}{2}-1} \exp\left(-\frac{y}{2}\right) \tag{3.7.12}$$

で与えられる。

ただし$\Gamma(\lambda) = \int_0^\infty x^{\lambda-1} e^{-x} dx$ (3.7.13)

自由度 N の χ ^{2}分布に従う確率変数 Yの期待値と分散は

$$\langle Y \rangle = \int_0^\infty y\mathrm{C}(y, N) dy = \frac{2}{\Gamma\left(\frac{N}{2}\right)} \int_0^\infty x^{\frac{N}{2}} \exp(-x) dx = \frac{2\Gamma\left(\frac{N}{2}+1\right)}{\Gamma\left(\frac{N}{2}\right)} = N \tag{3.7.14}$$

$$\langle Y^2 \rangle = \int_0^\infty y^2\mathrm{C}(y, N) dy = \frac{4}{\Gamma\left(\frac{N}{2}\right)} \int_0^\infty x^{\frac{N}{2}+1} \exp(-x) dx = \frac{4\Gamma\left(\frac{N}{2}+2\right)}{\Gamma\left(\frac{N}{2}\right)} = N(N+2) \tag{3.7.15}$$

$$\langle Y^2 \rangle - \langle Y \rangle^2 = 2N \tag{3.7.16}$$

そこで、reduced χ ^{2}

$$\tilde{S} \equiv \frac{S}{N-m} \tag{3.7.17}$$

を定義すると、

$\langle \tilde{S} \rangle = 1$ (3.7.18)

$\langle \tilde{S}^2 \rangle - \langle \tilde{S} \rangle^2 = \frac{2}{N-m}$ (3.7.19)

このとき、信頼度αでの reduced χ^2 の値、つまり

$\int_{S_1}^{\infty} C(yN, N)\, dy = 1 - \frac{\alpha}{2}$ (3.7.20)

$\int_{0}^{S_2} C(yN, N)\, dy = 1 - \frac{\alpha}{2}$ (3.7.21)

となる S_1, S_2 の値は下表のとおりである

自由度　N	信頼度　α=90%		信頼度　α=99%	
	S_1	S_2	S_1	S_2
1	0	3.84	0	7.88
5	0.23	2.20	0.08	3.35
10	0.39	1.83	0.22	2.52
50	0.69	1.44	0.54	1.59
100	0.78	1.24	0.67	1.40

直感的には、reduced X² をとると 1 近辺に分布する。1 から離れすぎているモデルは合わないということ。

例：　モデル関数を一次関数 $f(t,a,b) = at + b$ とする。

$$S(a,b) = \sum \frac{1}{{\sigma_i}^2}\left(x^{(i)} - at^{(i)} - b\right)^2$$

これが最小になる点は

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \frac{1}{D}\begin{pmatrix} \sum \frac{1}{{\sigma_i}^2} & -\sum \frac{t^{(i)}}{{\sigma_i}^2} \\ -\sum \frac{t^{(i)}}{{\sigma_i}^2} - & \sum \frac{\left(t^{(i)}\right)^2}{{\sigma_i}^2} \end{pmatrix}\begin{pmatrix} \sum x^{(i)}\, t^{(i)} \\ \sum x^{(i)} \end{pmatrix}, D = \left(\sum \frac{\left(t^{(i)}\right)^2}{{\sigma_i}^2}\right)\left(\sum \frac{1}{{\sigma_i}^2}\right) - \left(\sum \frac{\left(t^{(i)}\right)}{{\sigma_i}^2}\right)^2$$

いま上記の点のまわりで S をテーラー展開して

$$\left.\frac{\partial S(a,b)}{\partial a}\right|_{(\hat{a},\hat{b})} = 0, \left.\frac{\partial S(a,b)}{\partial b}\right|_{(\hat{a},\hat{b})} = 0$$

であることに注意すると

$$S(a,b) \cong S(\hat{a},\hat{b}) + \frac{(a-\hat{a})^2}{2}\left.\frac{\partial^2 S(a,b)}{\partial a^2}\right|_{(\hat{a},\hat{b})} + \frac{\left(b-\hat{b}\right)^2}{2}\left.\frac{\partial^2 S(a,b)}{\partial b^2}\right|_{(\hat{a},\hat{b})} + (a-\hat{a})\left(b-\hat{b}\right)\left.\frac{\partial^2 S(a,b)}{\partial a \partial b}\right|_{(\hat{a},\hat{b})}$$

$$S(a,b) \cong S(\hat{a},\hat{b}) + (a - \hat{a} \quad b - \hat{b})V\begin{pmatrix} a-\hat{a} \\ b-\hat{b} \end{pmatrix}$$

$$V=\begin{pmatrix} \left.\frac{1}{2}\frac{\partial^2 S(a,b)}{\partial a^2}\right|_{(\hat{a},\hat{b})} & \left.\frac{1}{2}\frac{\partial^2 S(a,b)}{\partial a\partial b}\right|_{(\hat{a},\hat{b})} \\ \left.\frac{1}{2}\frac{\partial^2 S(a,b)}{\partial a\partial b}\right|_{(\hat{a},\hat{b})} & \left.\frac{1}{2}\frac{\partial^2 S(a,b)}{\partial b^2}\right|_{(\hat{a},\hat{b})} \end{pmatrix}=\begin{pmatrix} \sum\frac{\left(t^{(i)}\right)^2}{{\sigma_i}^2} & \sum\frac{t^{(i)}}{{\sigma_i}^2} \\ \sum\frac{t^{(i)}}{{\sigma_i}^2} & \sum\frac{1}{{\sigma_i}^2} \end{pmatrix}$$

この逆行列が、誤差行列

$$\Sigma=\begin{pmatrix} {\sigma_a}^2 & c_{ab} \\ c_{ab} & {\sigma_b}^2 \end{pmatrix}=\frac{1}{\mathrm{D}}\begin{pmatrix} \sum\frac{1}{{\sigma_i}^2} & -\sum\frac{t^{(i)}}{{\sigma_i}^2} \\ -\sum\frac{t^{(i)}}{{\sigma_i}^2} & \sum\frac{\left(t^{(i)}\right)^2}{{\sigma_i}^2} \end{pmatrix}$$

となる。

注意：各測定点の誤差が既知でないときも、最小自乗法は使用できる。但し、誤差行列は、

$$\Sigma=\begin{pmatrix} {\sigma_a}^2 & c_{ab} \\ c_{ab} & {\sigma_b}^2 \end{pmatrix}=\frac{1}{\mathrm{D}}\begin{pmatrix} \sum\frac{1}{{\sigma_i}^2} & -\sum\frac{t^{(i)}}{{\sigma_i}^2} \\ -\sum\frac{t^{(i)}}{{\sigma_i}^2} & \sum\frac{\left(t^{(i)}\right)^2}{{\sigma_i}^2} \end{pmatrix}\frac{S\left(\hat{a},\hat{b}\right)}{N-m}$$

となる。

<<第 5 講ここまで>>

4 検定

<<第 6 講>>

4-1 仮説と検定

母集団分布について前もってなされる仮定を**統計的仮説**(“hypothesis”)という。
統計的仮説が正しいかどうか調べる作業を統計的検定（”test of hypothesis”）という。
検定とは、抽出された N ケの標本の値　$x^{(i)}: i = 1 \dots N$　を基にして、**仮説 H_0 を棄却するかどうかを判断する作業。**

同一の確率分布に従う N 個の確率変数$X^{(i)}: i = 1 \dots N$から決まる検定統計量（”test statictics”）
$K(X^{(1)}, \dots . X^{(N)})$を考え、仮説 H_0 を$H_0: \langle K \rangle = k_0$と記す
仮説 H_0 が成り立つと仮定したときの検定統計量の確率分布を H_0 を$W(k: H_0)$として、

$$\int_{R_0} W(k: H_0) dk = \alpha \tag{4.1.1}$$

となるように領域 R_0 を決める。この時αを有意水準(“significance level”)または危険率という。
もし、N ケの標本の値$x^{(i)}: i = 1 \dots N$から得られた値$K(x^{(1)}, \dots . x^{(N)})$が領域 R_0 に入ったら、
「危険率αで仮説 H_0 は棄却された」
という。ここで危険率とは、
「仮説 H_0 が正しいにもかかわらず、棄却されてしまう確率」
のことである。
〇片側検定：危険率αの危険域を領域の片側にのみ設定する方法。
〇両側検定：危険域を領域の両側に設定する検定法。危険率が合計でαとなるようにする。

4-2 正規分布に従う母集団に対する検定

i)　　母分散σ^{2}が既知であるときの母平均μに対する検定
仮説$H_0: \mu = \mu_0$　を信頼水準$1 - \alpha$で検定するには、

$$\bar{Z} \equiv \frac{\bar{X} - \mu_0}{\sigma / \sqrt{N}} \tag{4.2.1}$$

を検定統計量として両側検定で考える。$\bar{Z}$は正規分布$N(0,1)$に従うので、

$$P(z_0 \leq \bar{Z} \leq \infty) = \int_{z_0}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) dz = \frac{\alpha}{2} \tag{4.2.2}$$

となるz_0を求める。
例）N=10 の時α=0.05 とすると$z_0 = 1.96$である。
$\bar{Z}(x^{(1)}, \dots . x^{(N)}) \geq z_0$又は$\bar{Z}(x^{(1)}, \dots . x^{(N)}) \leq -z_0$ならば、仮説 H_0 は棄却される。

ii)　　母分散に対する検定

仮説$H_0: \sigma = \sigma_0$　を信頼水準$1 - \alpha$で検定するには、不偏分散　$S^2 = \frac{1}{N-1} \sum \left(X^{(i)} - \bar{X} \right)^2$を使って

$$\chi^2 \equiv (N-1) \frac{S^2}{\sigma^2} = \sum \frac{\left(X^{(i)} - \bar{X} \right)^2}{\sigma^2} \tag{4.2.5}$$

を検定統計量として両側検定で考える。これは、自由度 N- 1 のカイ 2 乗分布に従うので、

$$P(k_u \le \chi^2 \le \infty) = \frac{\alpha}{2} \qquad (4.2.6)$$

および

$$P(0 \le \chi^2 \le k_l) = \frac{\alpha}{2} \qquad (4.2.7)$$

となるk_l, k_uがχ^2の棄却域となる。

例題）測定数が 11 のデータに対して、σ を信頼度 α =0.1 で分散を検定する。この時のχ^2の棄却域を求めよ。
$\chi^2/10$の棄却域は 0.39,　1.83 より
χ^2に関して、$k_l = 3.90, k_u = 18.3$が棄却域となる。

iii)　母分散 σ ^{2}が未知であるときの母平均 μ に対する検定

仮説$H_0: \mu = \mu_0$　を信頼水準 1- α で検定するには、不偏分散　$S^2 = \frac{1}{N-1}\sum\left(X^{(i)} - \bar{X}\right)^2$を使って

$$T \equiv \frac{\bar{X} - \mu_0}{S/\sqrt{N}} \qquad (4.2.3)$$

を検定統計量として両側検定で考える。S≅ σ とできるときには i)と同じく正規分布を用いることができる。この時、T は自由度 N-1 の student 分布に従う。

$$P(t_0 \le T \le \infty) = \frac{\alpha}{2} \qquad (4.2.4)$$

となるt_0を求める。
例）α =0.05 とすると$t_0 = 2.26$である。
$T(x^{(1)}, \dots . x^{(N)}) \ge t_0$又は$T(x^{(1)}, \dots . x^{(N)}) \le -t_0$ならば、仮説 H_0は棄却される。

4-3 事前確率と事後確率
次の問題を考えてみよう
問）
狂牛病の検査試験の感度が
　狂牛病に罹っている場合 …　　　100.0%　陽性
　狂牛病に罹っていない場合 …　　99.8%　陰性、0.2%陽性
であるとする。あなたの飼っている牛 1 頭の検査を行ったところ、陽性反応が出た。
その牛が狂牛病に罹っている可能性は何%か。
答）
陽性反応が出たとしても、狂牛病に罹っていない可能性がある。

狂牛病に罹っている場合、陽性が出る可能性は 100%、
狂牛病に罹っていない場合、陽性が出る可能性は 0.2%。

よって、狂牛病に罹っている可能性は $\frac{100}{100+0.2} = 99.8\%$ であろう。

ところが、注意深く考えてみると、今の議論では、「検査前の時点における狂牛病に罹っている確率と罹っていない確率が 1 対 1 である」ことが暗黙のうちに仮定されている。
例えば、全国調査で、狂牛病に罹っている牛の割合が、10 万頭に対して 1 頭であることが分かっていた場合、陽性反応が出た牛が狂牛病に罹っている確率は

$$\frac{1\times100}{(1\times100)+(99999\times0.2)} \cong 0.5\%$$

である。
検査以前の確率分布（事前確率）が、検査後の確率分布（事後確率）に影響する。事前確率を仮定しなければ事後確率を決定することができない。

4-4 ベイズ原理（Bayes' Principle）

観測前の確率分布（事前確率）が、観測後の確率分布（事後確率）に影響する。事前確率を仮定しなければ事後確率を決定することができない。

以下のような記法で議論を進める
確率変数（random variable）：*X, Y, Z*, ...
標本値（sample value)または観測値（observation value）：*x, y, z*,
真の値、母数（parameter）：μ，θ，
確率密度関数（Probability Density Function: p.d.f.）: $f(x), f(\mu)$
条件付き確率密度関数（Conditional p.d.f.）$f(x|\mu), f(\mu|x)$

x, μの同時密度関数（joint density function）は (1.1.10)式より

$$f(x,\mu) = f(x|\mu)f(\mu) = f(\mu|x)f(x) \tag{4.4.1}$$

と書ける。これから、

$$f(\mu|x) = \frac{f(x|\mu)f(\mu)}{f(x)} = \frac{f(x|\mu)f(\mu)}{\int f(x|\mu')f(\mu')d\mu'} \tag{4.4.2}$$

この式により、ここで、$f(\mu)$を事前確率（prior probability)とすると、事前確率$f(\mu)$、観測された確率$f(x|\mu)$から事後確率(posterior probability) $f(\mu|x)$ が得られる。分母の値は規格化の為に必要。

- Bayes 原理を適用するには、事前確率が必要である。(Bayes 原理以外の推定方法では、事前確率は暗黙のうちに仮定されている)
- 事前確率の決め方については統一的見解はない。(統計学は事前確率の決定法については何も教えてくれない)

例題　6-1）　ポアソン分布
新種の素粒子の測定に成功した。観測数は 5 事象で、バックグラウンドイベントは無視できる程度であることが分かった。この素粒子が観測される確率が、生成断面積 1 fb あたり 5 事象となる実験条件であった時、生成断面積を統計誤差付きで述べよ。
答）方針：観測された事象数 x から観測事象数の期待値 μ を推定する問題。

○第 0 近似解：観測された事象数は Poisson 分布に従うので、
生成個数の期待値は$5 \pm \sqrt{5} = 5.0 \pm 2.2$個
よって、生成断面積は　(5.0±2.2)/5=1.0±0.4 fb である。
一見よさそうだが、推定範囲を 1 σ ではなく 3 σ にしたら生成断面積は　1.0±1.2 fb である。となり、あり得ない負の生成断面積まで推定範囲に入ることになる。

○より正確な解：ポアソン分布をきちんと考える。生成断面積の誤差を求められているので、αは片側α/2　ずつ考える。

$$\int_{\mu_u}^{\infty} \frac{y^x}{x!} e^{-y} dy = 1 - \alpha/2 \tag{4.4.3}$$

$$\int_{0}^{\mu_l} \frac{y^x}{x!} e^{-y} dy = 1 - \alpha/2 \tag{4.4.4}$$

として、上限(μ_u)、下限(μ_l)を考える。ここで、

$$\int_0^{\infty} \frac{\mu^x}{x!} e^{-\mu} d\mu = \frac{(-1)^x}{x!} \left[\frac{d^x}{da^x} \int_0^{\infty} e^{-a\mu} d\mu\right]_{a=1} = \frac{(-1)^x}{x!} \left[\frac{d^x}{da^x} \frac{1}{a}\right]_{a=1} = 1 \tag{4.4.5}$$

を使った。

上限(μ_u)、下限(μ_l)に関して、手計算はできないので、表から数字を拾う。（付表 3）
信頼度 90%(1-α=0.9)で観測事象数の期待値は　$\mu_l = 1.8, \mu_u = 9.9$ より　$\mu = 5^{+4.9}_{-3.2}$　である。
したがって、生成断面積は$1^{+1.0}_{-0.6}$　fb となる。
信頼度 95%(1-α=0.95)で観測事象数の期待値は　$\mu_l = 1.8, \mu_u = 11.3$ より $\mu = 5^{+6.3}_{-3.2}$ である。
したがって、生成断面積は$1^{+1.3}_{-0.6}$　fb となる。

○事前確率の考えを取り入れた解。観測された事象数 x に対して生成個数の期待値 μ を推定しているのだが、μ は 0 以上であることを考慮に入れ、事前確率

$$f(\mu) = \begin{cases} 0 & : \mu < 0 \\ 1 & : \mu \geq 0 \end{cases}$$

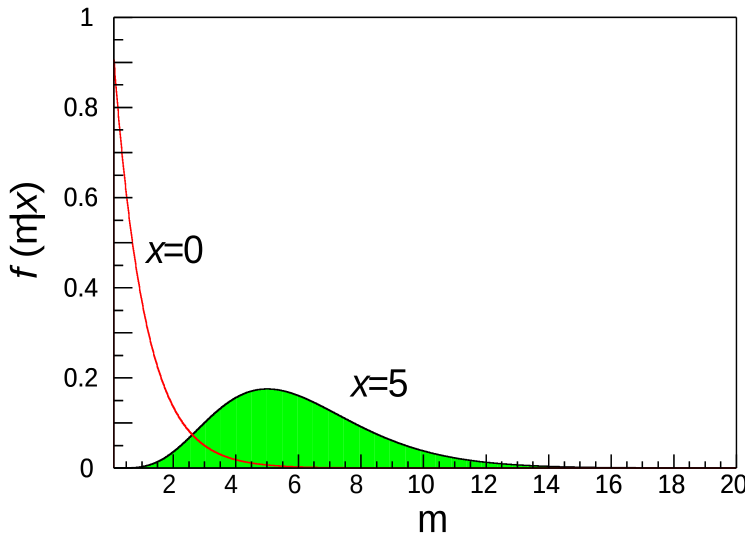
を導入する。ベイズの原理から

$$f(\mu|x) = \frac{f(x|\mu)f(\mu)}{f(x)} = \frac{f(x|\mu)f(\mu)}{\int_0^{\infty} f(x|\mu')f(\mu')d\mu'} = \frac{f(x|\mu)}{\int_0^{\infty} f(x|\mu')d\mu'} = f(x|\mu) = \frac{\mu^x}{x!} e^{-\mu} \tag{4.4.6}$$

$$E(f(\mu|x)) = \int_0^{\infty} \mu\, f(\mu|x) d\mu = \int_0^{\infty} \mu \frac{\mu^x}{x!} e^{-\mu} d\mu = -\left[\frac{\mu^{x+1}}{x!} e^{-\mu}\right]_0^{\infty} + (x+1) \int_0^{\infty} \frac{\mu^x}{x!} e^{-\mu} d\mu$$

$$= x + 1\int_0^\infty f(\mu|x)\mathrm{d}\mu = x + 1 \tag{4.4.7}$$

よって、期待値は x+1=6 事象となるようなポアソン分布を考える。
0 次近似では生成断面積は　$6 \pm \sqrt{6} = 6.0 \pm 2.4$ fb　である。生成断面積の最頻値は 5.0 fb である。
ポアソン分布の表から、
信頼度 90%で　観測事象数の期待値は$\mu_l = 2.2, \mu_u = 11.5$　$6.0^{+5.5}_{-2.8}$ である。
したがって、生成断面積は$1.2^{+1.1}_{-0.6}$　fb となる。
信頼度 95%で　観測事象数の期待値は$\mu_l = 2.2, \mu_u = 12.8$　$6.0^{+6.8}_{-2.8}$ である。
したがって、生成断面積は$1.2^{+1.4}_{-0.6}$　fb となる。



例題　6-2）正規分布

放射線源Ａからのγ線の、1 分間での検出数測定を 100 回行った結果、

その(標本)平均 $\bar{x} = \frac{1}{100}\sum_{i=1}^{100} x_i$ は 24.0

その(標本)不偏分散 $S^2 = \frac{1}{99}\sum_{i=1}^{100}\left(x_i - \bar{x}\right)^2$ は 25.0

であった。

(1)　γ線の検出効率（検出効率＝検出されたγ線の個数÷放出されたγ線の総数）が 10%であるとして、放射線源Ａの強度 I（Bq：1 秒間に放出するγ線の個数）の範囲を 95%の　信頼度で推定しなさい。（有効数字 3 桁）但し、測定結果の平均値は正規分布に従い、その分散は標本不偏分散と等しいとする。

(2)　放射線源Ａと同じエネルギーのγ線を放出する、別の放射線源Ｂがある。上記と同様の測定条件で測定したところ(つまり、検出効率は 10%で等しい)、10 分間の計数（検出したγ線の個数）が０であった。放射線源Ｂの強度の上限値を 95%の信頼度で推定しなさい。（有効数字 2 桁）

解答例

(1) $\mathrm{Y} = 24.0 \div 0.10 \div 60 = 4.00$

(3.4.1)より $\sqrt{\sigma_Y^2} = \sqrt{\frac{25}{100}} \div (0.10 \times 60) = \frac{5}{60} = 0.0833$

両側 95%の信頼度なので、正規分布の表より 1.96σをとる。

α	u
0.025	1.96

$$Y - 1.96 \times \sigma_Z < I < Y + 1.96 \times \sigma_Z$$
$$3.83 < I < 4.16$$

(2)

$$P(0) = e^{-\mu'} \geq 0.05$$
$$\therefore \mu' \leq -\ln 0.05 = \ln 20 = 2\ln 2 + \ln 5 = 3.0$$
$$I' \leq \frac{3.0}{0.10 \times 10 \times 60} = 0.0500$$

(3.0 は表から拾ってもよい。)

例題　6-3）正規分布
ある年の７月後半１４日間の最高気温を東京、名古屋、仙台は次の表のようになった。

	仙台	東京	名古屋
標本平均	24.77	29.71	31.22
不偏分散	7.39	7.10	7.74

仮説「仮説 H_0：名古屋は仙台より暑い」と「仮説 $H_0{}^C$：東京と仙台の平均気温が等しい」を検定せよ。

解答例
「仮説 $H_0{}^C$：名古屋と仙台の平均気温が等しい」が棄却できるか検定する。仙台、名古屋の最高気温をそれぞれ確率変数 X, Y とする。$K=X-Y$ とおくと、「仮説 $H_0{}^C$：$\langle K \rangle = 0$」である。そこで、(4.2.3)などより

$$S = \sqrt{\frac{S_X^2}{14} + \frac{S_Y^2}{14}} = \frac{\sqrt{7.39^2 + 7.74^2}}{\sqrt{14}} = 2.86$$

$$T = \frac{\bar{X} - \bar{Y}}{S} = \frac{24.77 - 31.22}{2.86} = -2.26$$

を検定統計量として片側検定で考える。自由度 26(14×2-2)の student 分布に従うのであるから、危険率 α=0.05 とすると $t_0 = 2.06$ である。$T(x^{(1)}, \ldots . x^{(N)}) \leq -t_0$ なので、仮説「仮説 $H_0{}^C$：名古屋と仙台の平均気温が等しい」は棄却され、「仮説 H_0：名古屋は仙台より暑い」が正しい。
同様に、東京と仙台を比べると、

$$T' = \frac{\bar{X} - \bar{Z}}{S} = \frac{24.77 - 29.71}{2.74} = -1.80, S = \sqrt{\frac{S_X^2}{14} + \frac{S_Y^2}{14}} = \frac{\sqrt{7.39^2 + 7.10^2}}{\sqrt{14}} = 2.74$$

となり、仮説「仮説 $H_0{}^C$：東京と仙台の平均気温が等しい」は棄却できない。

物理実験学Ｉ（統計）講義に関する意見と改善案(2012 年度)

＊もっと例題がほしかった。式を見るよりは実際に問題を解く方が理解しやすいと思う。
＊式だけではイメージつかみにくいので、実際の問題を。講義で例題を解いてほしい。
⇒講義中に扱える例題を増やします。講義ノートに例題を追加しておきます。

物理実験学 I （統計）まとめ地図

<基礎概念>

期待値　$E(X) = \int x \cdot f(x)dx$

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

<確率分布>

ポアソン分布　$P(x,\mu) \equiv \frac{\mu^x}{x!}\mathrm{e}^{-\mu}$　期待値= μ　分散= μ

正規分布　$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})$　期待値= μ　分散= σ^2

<使用例>

* μ を知っている時にある測定値 x を得る確率：確率分布関数から計算

性質

大数の法則　$P(|\bar{X} - \mu| < \varepsilon) \leq 1 - \frac{\sigma^2}{N\varepsilon^2}$:　測定回数を重ねると平均値は真の期待値に近づく。

中心極限定理：　N 回の試行で、$\bar{X}$は　期待値 μ 、分散$\frac{\sigma^2}{N}$の正規分布に近づく

<母数推定>

最尤法：　$(\theta) \equiv W\left(x^{(1)},\theta\right)\mathrm{W}(x^{(2)},\theta) \ldots \mathrm{W}\left(x^{(N)},\theta\right) = \prod W\left(x^{(i)},\theta\right)$　を最大にする。

最小二乗法：$S = \sum \frac{1}{{\sigma_i}^2}\left(x^{(i)} - f\left(t^{(i)},\theta\right)\right)^2$　を最小にする

S は自由度 $N-m$ の χ^2分布に従う

<最終目的：母数推定・検定>

標本平均　$\bar{X} = \frac{1}{\mathrm{N}}\sum_i X^{(i)}$　$\Rightarrow$　$\mu \sim \bar{X}$

不偏分散　$S^2 \equiv \frac{N}{N-1}V^2 = \frac{1}{N-1}\sum\left(X^{(i)} - \bar{X}\right)^2$　$\Rightarrow$　$\sigma \sim S/\sqrt{N}$

信頼度　σ 既知（もしくは S で代用）　μ を求める　$\Rightarrow$　正規分布

σ を求める　$\Rightarrow \chi^2$ 分布

物理実験学 I （統計）付表

Table 33.1: Area of the tails α outside $\pm\delta$ from the mean of a Gaussian distribution.

α	δ	α	δ
0.3173	1σ	0.2	1.28σ
4.55×10^{-2}	2σ	0.1	1.64σ
2.7×10^{-3}	3σ	0.05	1.96σ
6.3×10^{-5}	4σ	0.01	2.58σ
5.7×10^{-7}	5σ	0.001	3.29σ
2.0×10^{-9}	6σ	10^{-4}	3.89σ

付表１　母分散 σ^2 の正規分布で信頼度 $1-\alpha$ を与える δ。

2011 Review of Particle Physics　(J. Phys. G **37**, 075021 (2010))

http://pdg.lbl.gov/2011/reviews/contents_sports.html

Table 33.3: Lower and upper (one-sided) limits for the mean ν of a Poisson variable given n observed events in the absence of background, for confidence levels of 90% and 95%.

	$1-\alpha=90\%$		$1-\alpha=95\%$	
n	$\nu_{\rm lo}$	$\nu_{\rm up}$	$\nu_{\rm lo}$	$\nu_{\rm up}$
0	–	2.30	–	3.00
1	0.105	3.89	0.051	4.74
2	0.532	5.32	0.355	6.30
3	1.10	6.68	0.818	7.75
4	1.74	7.99	1.37	9.15
5	2.43	9.27	1.97	10.51
6	3.15	10.53	2.61	11.84
7	3.89	11.77	3.29	13.15
8	4.66	12.99	3.98	14.43
9	5.43	14.21	4.70	15.71
10	6.22	15.41	5.43	16.96

付表２　n 事象観測された際のポアソン分布の期待値 ν の上限値もしくは下限値。

2011 Review of Particle Physics　(J. Phys. G **37**, 075021 (2010))

http://pdg.lbl.gov/2011/reviews/contents_sports.html

Table 33.4: Unified confidence intervals $[\nu_1, \nu_2]$ for a the mean of a Poisson variable given n observed events in the absence of background, for confidence levels of 90% and 95%.

	$1-\alpha = 90\%$		$1-\alpha = 95\%$	
n	ν_1	ν_2	ν_1	ν_2
0	0.00	2.44	0.00	3.09
1	0.11	4.36	0.05	5.14
2	0.53	5.91	0.36	6.72
3	1.10	7.42	0.82	8.25
4	1.47	8.60	1.37	9.76
5	1.84	9.99	1.84	11.26
6	2.21	11.47	2.21	12.75
7	3.56	12.53	2.58	13.81
8	3.96	13.99	2.94	15.29
9	4.36	15.30	4.36	16.77
10	5.50	16.50	4.75	17.82

付表 3　n 事象観測された際のポアソン分布の期待値 ν の上限値および下限値。

2011 Review of Particle Physics　（J. Phys. G **37**, 075021 (2010)）

http://pdg.lbl.gov/2011/reviews/contents_sports.html