工学へのシステム信頼性

目次

1 信頼性、安全性、リスク

1.1 信頼性工学

対象とするいくつかの部品から構成されているシステムがもっている機能について、状況を把握できないある原因から働かなくなってしまった状態を故障 (failure) という。信頼性が高いということは、故障となるような状態に陥らないことであるが、複雑に絡み合った要素からは簡単に故障を引き起こすことを少なるすることは簡単ではない。また故障となる要因をもつ部品に対して、信頼性が高くなるようにするにはどうしたらよいか？信頼性を高めることが果たしてできることであろうか？

システム規模の巨大化にともない、構成が極めて複雑で、機能がより高めた要素がつくる高度化されたシステムには、システムの故障が生じると、社会的、経済的な損失のみならず、人命や財産あるいは環境の破壊など影響は極めて大きくなる。単純なものからでは計り知れないことが起こり得る。

高度化されたシステムの信頼性の考察には、故障の発生があってもシステム全体が機能を保ちつつ、影響の減少を図りつつ、正常に稼働できるよう工夫や保全が考えられていなければならない。そのために、故障発生の定量化とシステム機能の保全維持には十分か解析と、それらの関連付けを定めておくことことが必要となるであろう。

ここではこれら信頼性の高度化と保全維持とを目的とし、その定量化に必要な数学的手段、さらには確率、統計データの分析手法をもちい、コンピュータ支援の計算解析を取り入れて、システムの目的を効率化するような数理的最適化理論あるいは変分問題の解析を取り入れながら、信頼性の尺度として、信頼度を定義します。日本工業規格にはつぎのように定められている。

定義 1.1 信頼度 (riliability) の定義 (JIS Z 8115) 「規定の条件のもとで、意図する期間中、規定の機能を遂行する確率」

システムには (i) 機能、 (ii) 使用の条件 および (iii) 使用期間という 3 つの定められた規律がありますから、その故障の定量化には、システムの故障（破損）の状態を明確に定義しなければなりません。のちに使用条件、使用の環境を定め、期間を指定したうえで、意図した機能を達成する確率などを計算します。このような信頼度の算定する理論や分析のための手法を一般に信頼性理論 (reliability theory) とよぶことにします。

1.2 不確定な要因

いま信頼度 R が与えられてとき、故障確率 (不信頼度の尺度)P_f を

$$P_f = 1 - R$$

と定める。故障によって生じる被害の大きさ D が分れば、リスク (Risk) をつぎで定量化する。

$$\mathrm{Risk} = D \times P_f = D \times (1 - R)$$

システムが課せられた機能する維持し継続することは、システムが故障せず、もし故障しても修理を迅速に行うことができれば、そのシステムの目的を達成している。端的にいえば、「故障しにくく」、「故障しても容易に直せる」ことができればシステムは機能を達成している。つまり、故障しにくい「信頼性」と直しやすさの「保全性」が重要な概念である。保全性 (maintainability) の尺度として

定義 1.2 保全度の定義 「故障したシステムの修理が規定期間内に完了する確率」

さらに、信頼度と保全度を組み合わせた評価尺度、機能有効性 (availability) として、

定義 1.3 機能有効性の定義 「ある使用条件のもとで特定の時期にシステムが所定の機能を発揮している確率」

平均故障時間間隔 (MTBF, Mean Time between Failures) 平均修理時間 (MTTR, Mean Time to Repair) をもちいて

$$\mathrm{Availability} = \frac{MTBF}{MTBF + MTTR}$$

信頼性工学 (reliability engineering) では、信頼度の評価を正確に計算するだけでなく、システムのライフサイクルに必要な費用を最小化すること、使用するシステムの効率を高めるような、設計、製造、運用、管理を考える総合的な工学の分野です。

✍ ✍ ✍ **演習問題** ✍ ✍ ✍

2 確率統計の基礎数理

数理的な分析には、変動を表現するための確率統計の知識を復習しておきましょう。

2.1 確率論の基礎

データを集計整理すると、数値的な集約すること、すなわち四則演算やデータから関数を取り扱うためには、数値の結果を、偶然変動する数学的な方法として、確率を導入する。数学の「公理」としていくつかの性質を満たす概念によって定義される。加法性測度論であり、この考え方は、ロシアの数学者, A.N.Kolmogorov（コルモゴロフ, 1903-1987）により、大きな進展をもたらされた。彼以前の確率論はラプラスによる「確率の解析的理論」に基づく古典的確率論が中心であったが、彼が「測度論に基づく確率論」「確率論の基礎概念（1933 年）」で公理主義的確率論を立脚させ、現代確率論の始まりとなった。(Wikipedia より)

事象 A, B, A の否定、補事象；$\overline{A}$、 和事象 $A \cup B$, 積事象 $AB = A \cap B$

確率 $P(A)$, 加法性；$P(A \cup B) = P(A) + P(B) - P(AB)$

確率変数 (random variable, r.v.); X, Y など変動の結果を表す変数

確率密度関数 (probability density function, pdf); 変数に対して起こり得る確率の値。$f_X(x), (x > 0, -\infty < x < \infty)$, $p_X(k) (k = 0, 1, 2, \cdots)$、関数値として表現式や確率の数値を示す。

確率分布関数 (probability distribution function, DF)；取り得る範囲、区間に対して定めた確率の値, ゼロ（起こりえない値）から 1（当然の結果としてみなされる値）までの範囲。

期待値 (expectation) 連続型 $E[X] = \displaystyle\int_{-\infty}^{\infty} f_X(x)dx$, 離散型 $E[Y] = \displaystyle\sum_{k=0}^{\infty} p_Y(k)$ 、

２次モーメント $E[X^2] = \displaystyle\int_{-\infty}^{\infty} x^2 f_X(x)\,dx$, $E[Y^2] = \displaystyle\sum_{k=0}^{\infty} k^2\,p_Y(k)$、

$$E[Y(Y-1)] = E[Y^2] - E[Y] = \sum_{k=0}^{\infty} k(k-1)\,p_Y(k) = \sum_{k=0}^{\infty} k^2\,p_Y(k) - \sum_{k=0}^{\infty} k\,p_Y(k)$$

平均（期待値） (分布の平均)(mean) $\mu_X = E[X]$, 「平均値」はデータの算術平均を表す。期待値は重みつき平均で求められ、和だけではなく、積分の場合で計算されることもある。

標準偏差 (standard deviation, STD),

分散 (variation)、 $\sigma_X^2 = V[X] = E[X^2] - (\mu_X)^2$, 母集団分散、標本分散

多次元確率変数 (multivariate random variable)

結合分布 、

周辺分布 、

条件つき分布 、

確率変数の独立性

線形一次結合 、

線形関数の期待値、分散

2.2 データ分析の基礎

データ変量 サイズ (標本の大きさ) n; $\{x_1, x_2, \cdots, x_n\}$

2 変量データ サイズ n; $\{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$

3 変量データ (ベクトル表示)　サイズ n; $\left\{ \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}, \cdots, \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} \right\}$

標本平均　$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\overline{y} = \frac{\sum_i y_i}{n} = \sum_i \frac{y_i}{n}$,

標本分散　$s_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$

標本不偏分散　$u^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n(n-1)/2}\sum_{i<j}\frac{(x_i - x_j)^2}{2}$,

例. $n = 3$ ならば　$s_x^2 = \frac{1}{3}\left(\frac{(x_1 - x_2)^2}{2} + \frac{(x_1 - x_3)^2}{2} + \frac{(x_2 - x_3)^2}{2}\right)$

$= \frac{1}{3}(x_1^2 + x_2^2 + x_3^2) - \left(\frac{x_1 + x_2 + x_3}{3}\right)^2$

[注意]： $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$、各データと平均値の差 n 個の総和はゼロとなる。「データ相互間の $1 \le i < j \le n$ とした 2 乗偏差 $(X_i - X_j)^2$ の組合わせ個数 ${}_nC_2$ 個の標本平均、変形すると $n-1$ で割るものと等しい」

$$\begin{aligned} u^2 &= \frac{2}{n(n-1)}\left\{(X_1 - X_2)^2 + (X_2 - X_3)^2 + \cdots + (X_{n-1} - X_n)^2\right\} \\ &= \frac{1}{n-1}\left\{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \cdots + (X_n - \overline{X})^2\right\}; \end{aligned}$$

標本標準偏差　$s = \sqrt{s^2}$, $s_y = \sqrt{s_y^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2}$、標本分散の平方根で、平均値と同じ単位をもち、「(平均値) プラス／マイナス (標準偏差)」という演算が意味をもつ 。$\overline{X} = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$; 「母平均は母集団の平均値（期待値）で、標本平均はデータの算術平均」; 標本分散 (sample variance); 「母分散は母集団の分散」、「標本分散は 2 通りの統計量があり、標本分散 (sample variance)s^2(sample の s の意味) と標本不偏分散 (sample unbiased variance)」、不偏 (unbiased) とは偏りがないという意味で、u^2 を使うことが多い。$s^2 = s_x^2 = \frac{1}{n}\left\{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \cdots + (X_n - \overline{X})^2\right\} = \frac{1}{n}(X_1^2 + X_2^2 + \cdots + X_n^2) - \overline{X}^2$;

変動係数　(coefficient of variation, CV) $\left(\mathrm{CV} = \frac{\sqrt{s_x^2}}{\overline{x}} = \frac{s_x}{\overline{x}}\right)$

標本共分散

標本相関係数　、

平均ベクトル　、

共分散行列

n-データ変量 (sample n data)：$X = (X_1, X_2, \cdots, X_n)$; 標本データは、母集団から無作為（ランダム）抽出した結果の数値であり、母集団分布を確率分布とする確率変数である。無作為（ランダム）という意味は、確率変数として、独立かつ同一分布にしたがうということで、各変数は $EX_i = \mu$, $V[X_i] = \sigma^2$, $E[X_i^2] = \sigma^2 + \mu^2$,

標本平均値 (sample average) $\overline{X}$ の平均（期待値）　：同一分布性から、$E[\overline{X}] = \frac{1}{n}\sum_i E[X_i] = \frac{1}{n}\sum_i \mu = \mu$

標本不偏分散 (unbiased variance)$u^2 = u_x^2$ の平均（期待値）　：独立性から $V[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n}\sigma^2 = n\sigma^2$,

$V[\overline{X}] = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n}\sigma^2$, $E[(X_i - X_j)^2] = 2\sigma^2$ この式から $E[u^2] = \sigma^2$;

標本標準偏差　; $SD = \sqrt{u^2}$;

変動係数 ; CV（Coefficient of Variation）($\sqrt{s_n^2}/\overline{X}$);「不偏的な統計量で母集団の比較にもちいる」

平均 2 乗偏差 ：m_2 、

平均絶対偏差 $E[|X-\mu_x|]$ 平均との差の絶対値に対する期待値

平均絶対偏差: 「平均と各データとの差を絶対値で平均する」; 標本共分散「各平均との偏差の積 $(X_i-\overline{X})(Y_i-\overline{Y})$ を平均する」、標本相関係数「標準化した各データの積 $\dfrac{(X_i-\overline{X})}{\sigma_X}\dfrac{(Y_i-\overline{Y})}{\sigma_Y}$ を平均する」、平均ベクトル「各データの平均値を並べたベクトル」、共分散行列「2 個のデータに対してつくった積の値を成分とする行列」

トリム平均値

中央値（メディアン）

最頻値（モード）

四分位範囲、第 k 四分位数

歪度（わいど、skewness）

尖度 (せんど, kurtosis)

ピアソンの相関係数

スピアマンの順位相関係数

単回帰と重回帰

偏相関

多重共線性

線形モデル式 2 個の説明変数 (x^1, x^2) から、y を表す式；

$$y = \beta_0 + \beta_1 x^1 + \beta_2 x^2 + \epsilon$$

ϵ は誤差変動

参照：

記述統計学:`http://www.math.s.chiba-u.ac.jp/~yasuda/statEN2/descriptiveA21.pdf`

統計学A: `http://www.math.s.chiba-u.ac.jp/~yasuda/statA0503.html`

統計学 B: `http://www.math.s.chiba-u.ac.jp/~yasuda/statB02.html`

確率統計の演習問題: `http://www.math.s.chiba-u.ac.jp/~yasuda/statEN2/toi4U.pdf`

3 シミュレーション

3.1 コイン投げとサイコロ

- 表計算ソフトによるコイン投げ、サイコロの結果：
  `http://www.math.s.chiba-u.ac.jp/~yasuda/statA0503.html`
- 統計でつかう表計算ソフトの解説、データ分析のアドイン：
  `http://www.math.s.chiba-u.ac.jp/~yasuda/statA0503.html`
  `http://www.math.s.chiba-u.ac.jp/~yasuda/statA/stat-excel-14.pdf`
- 統計処理のためのソフトウェア R：
  `http://www.math.s.chiba-u.ac.jp/~yasuda/shikou/Rdocmake/R-introY.pdf`

3.2 乱数の生成

乱数シミュレーションの資料：`http://www.math.s.chiba-u.ac.jp/~yasuda/sysKOU/simulation.pdf`

(i) 逆関数法（分布関数からの逆関数を用いる方法）
(ii) 棄却採択法 （一様乱数から任意の分布にしたがう乱数の発生法）
(iii) 正規乱数の生成 （中心極限定理による正規分布にしたがう乱数の生成法）

4 信頼性理論でよく表れる確率過程

時間パラメータを入れた状態変化を表現する確率過程を説明します。 (i) 2 項過程, (ii) ポアソン過程, (iii) マルコフ過程:
`http://www.math.s.chiba-u.ac.jp/~yasuda/sysKOU/kakuKatei.pdf`

4.1 2 項過程

時刻は離散時刻をパラメータとし、状態は「増える」と「増えない」の２通りで、窓口への客の来訪、来訪なしの状態で変化ししていきます。

4.2 ポアソン過程

連続時刻パラメータとし、窓口への客の来訪、来訪なしの変化は間隔期間を指数分布としたモデルを考えていきます。

4.3 マルコフ過程

より一般的な状態変化を、推移行列で表し、時刻変化も離散的なマルコフ連鎖のモデルや連続的（微小時刻で変化する）に動くブラウン運動やより一般的なマルコフ過程を考えます。

5 最適化法、線形計画法

最適化理論：`http://www.math.s.chiba-u.ac.jp/~yasuda/toku05.html`
線形計画法：`http://www.math.s.chiba-u.ac.jp/~yasuda/toku05/linearprogm02a.pdf`

6 複雑ネットワーク

相関構造の可視化
隣接行列、隣接リスト
R の igraph ライブラリー
矢久保孝介、「複雑ネットワークとその構造」共立出版　２０１３
Newman,M.E.J.; Networks:Introduction,Oxford University Press, 2010

7　機械学習の基礎概念

１９９０年以降インターネットの進化と計算機の処理能力の向上、記憶容量の増大などにより、いままでは処理されていなかった対象を取り込み、ビックデータとよばれる大規模なデータを処理対象とできる時代になってきている。大規模なデータから、必要とする求められるデータの統計的な処理を発掘する、データマイニングといういわゆる有用な役立つ情報を引き出す抽出方法での、数理的なモデルが「機械学習」とよばれる。このために（１）統計データを扱うために基礎となる数理統計学と（２）モデル構築をするための最適化理論、が重要な基礎理論となる。

- 確率分布のパラメータ推定、ベイズ統計学
- 予測の精度向上、バイアス・バリアンス分解
- 識別モデル学習、パターン認識（サポートベクトルマシーン, SVM）

8　信頼性の基本量

離散データから連続データへの極限

時刻 t まで故障した個数の比率 $F(t)$: t における故障比率（不信頼度）

時刻 t まで故障していない個数の比率 $R(t) = 1 - F(t)$: t における信頼度（残存確率）

故障の起きた時間、故障時間 (failure time）は確率変動するから確率変数で表される。その故障時間の密度関数

定義 8.1　瞬間故障率 (hazard rate)、たんに故障率 (failure rate) の定義 「時刻 t までに残存していたシステムの個数比率 $R(t)$ に対して、引き続き瞬間時間内（微小時間内）$\dfrac{d}{dt}F(t) = f'(t)$ に故障するシステムの個数比率 $f(t)$ の条件つき確率」

$$\lambda(t) = \frac{F'(t)}{R(t)} = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)}$$

8.1　故障率のパターン

8.2　故障時間の分布

定理 8.1　aaa aaa aaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaa aaa aaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaa aaa aaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaaaa 直角三角形の斜辺の長さを c とし、その他の辺の長さを a, b とした時 $a^2 + b^2 = c^2$ なる関係が成立する。

定理 8.2　`\begin{unshadedtheorem}` 直角三角形の斜辺の長さを c とし、その他の辺の長さを a, b とした時 $a^2 + b^2 = c^2$ なる関係が成立する。

8.3 信頼性の基本量

離散データから連続データへの極限

時刻 t まで故障した個数の比率 $F(t)$: t における故障比率（不信頼度）

時刻 t まで故障していない個数の比率 $R(t) = 1 - F(t)$: t における信頼度（残存確率）

故障の起きた時間、故障時間 (failure time）は確率変動するから確率変数で表される。その故障時間の密度関数

定義 8.2 瞬間故障率 (hazard rate)、たんに故障率 (failure rate) の定義 「時刻 t までに残存していたシステムの個数比率 $R(t)$ に対して、引き続き瞬間時間内（微小時間内）$\dfrac{d}{dt}F(t) = f'(t)$ に故障するシステムの個数比率 $f(t)$ の条件つき確率」

$$\lambda(t) = \frac{F'(t)}{R(t)} = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)}$$

8.4 故障率のパターン

8.5 故障時間の分布

定理 8.3 aaa aaa aaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaa aaa aaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaa aaa aaaa aaaaa aaaaa aaaaa aaaaa aaaaa aaaaa 直角三角形の斜辺の長さを c とし、その他の辺の長さを a, b とした時 $a^2 + b^2 = c^2$ なる関係が成立する。

定理 8.4 `\begin{unshadedtheorem}` 直角三角形の斜辺の長さを c とし、その他の辺の長さを a, b とした時 $a^2 + b^2 = c^2$ なる関係が成立する。

定理 8.5（ピタゴラス君とアルキメデスさん） 直角三角形の斜辺の長さを c とし、その他の辺の長さを a, b とした時 $a^2 + b^2 = c^2$ なる関係が成立する。

定義 8.3（名前のない定義） 直角三角形の斜辺の長さを c とし、その他の辺の長さを a, b とした時 $a^2 + b^2 = c^2$ なる関係が成立する。

✍ ✍ ✍ **演習問題** ✍ ✍ ✍

9 信頼性データの統計分析

◇ 記述統計 (descriptive statistics)

◇ 推測統計 (inferential statistics)

9.1　統計データの処理

9.2　確率分布の当てはめ

9.3　適合度の検定

✍ ✍ ✍　演習問題　✍ ✍ ✍

10　システムの信頼性

10.1　直列、並列システムの信頼性

10.2　一般システムの信頼性

10.3　構造関数

✍ ✍ ✍　演習問題　✍ ✍ ✍